DISSERTATION

RAND

What is the Quality of Care in a Developing Country?

Measuring Physician Practice and Health Outcomes

Jorge A. Muñoz

RAND Graduate School

20020805 174

DISSERTATION

RAND

What is the Quality of Care in a Developing Country?

Measuring Physician Practice and Health Outcomes

Jorge A. Muñoz

RGSD-163

RAND Graduate School

ACKNOWLEDGMENTS

TABLE OF CONTENTS FOR CHAPTER 1

RESEARCH MOTIVATION AND OBJECTIVES

The objectives of this dissertation are to provide a better understanding on how to measure the quality of health care in a less developed country in an effort to furnish an instrument for improving clinical care and national health policy formulation. The audience of this work thus includes health care practitioners as well as policymakers. To accomplish the stated objectives, this research will initially analyze the relationship between structural measures of quality with the clinical knowledge (process) of physicians. Subsequently, we will analyze the impact of clinical knowledge, and by association, clinical practice of primary health providers on the health status of a population.

Improving health care requires skillful practitioners to selectively use their knowledge, medical expertise, and medical technologies. The combination of what providers do and what they have to work with are the basic inputs that are expected to result in better health. When better health does not result, however, the responsibility is often assigned to the lack of material inputs. This is particularly true in developing countries where clinical facilities and medicine are obviously lacking. And while material inputs could be wanting, poor clinical practice could be an even more important explanation.

Data from the US and to a lesser extent from other countries has begun to reveal wide variation in clinical practice (Eddy 1984, Bhat 1999). And although some of the variation in medical care is attributable to regional practices, customs, and other environmental variables that do not adversely affect health outcomes (Brook 1984, Chassin 1987, Berwick 1991, Eddy 1984, Weiner 1995), there is certainly variability that is attributable to controllable extrinsic sources. Some of the controllable extrinsic sources of variation in medical practice are the training standards and the inadequate application of

medical knowledge and science (Kalf 1996, Roos 1994); both of these sources are quantifiable.

This dissertation will explore variations in the area of medical knowledge in a less developed country first, and then examine the relationship of the knowledge levels with structural aspects of health care, including resource availability. A final analysis will examine the relationship of knowledge levels with health outcomes at the population level. This will be done approaching the variations from a quality of medical care or "quality of care" perspective.

RESEARCH FRAMEWORK

The definition of quality of care according to the Institute of Medicine is "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" (Lohr 1990). According to Avedis Donabedian, the herald of the quality of care field, the definition of quality of care is "that kind of care which is expected to maximize an inclusive measure of patient welfare, after one has taken account of the balance of expected gains and losses that attend the process of care in all its parts." (Donabedian 1980). As a research framework, quality of care is defined as the "component of the difference between efficacy and effectiveness that can be attributed to care providers, taking account of the environment in which they work." (Brook & Lohr 1991).

The standard framework for measuring quality of health care provision is based on three measures of quality: structure, process, and outcome (Donabedian 1980, De Geynt 1995, McGlynn 1995, Brook 1996). The three measures are briefly described as follows:

- Structural Measures—physician characteristics and physical means (i.e. hospital, equipment)
- Process Measures—the factors of the encounter between the health care provider and the patient
- Outcome Measures—the subsequent health status of the patient.

STRUCTURAL MEASURES

This measure evaluates the characteristics of the health system, including the personnel and facilities that provide health care and how they are organized. The specialty training of physicians, their years

of practice, the accreditation of physicians and hospitals, as well as the ownership of the hospital are all components of this measure.

These measures are relatively inexpensive to obtain and are readily available (De Geynt 1995). However, structural quality is considered a necessary but not sufficient condition for better health outcomes (Nordyke 2000). For structural measures to be to be a valid dimension of quality, there must be evidence that variations in structural measures lead to changes in patients' outcomes (Brook 1996).

There is inconsistent and even paradoxical evidence on the effects of structural measures on process of care (Brook 1990). There is also a lack of empirical evidence of the effects of structural measures like hospital accreditation, on the variation of hospital quality in the US (Keeler 1992).

PROCESS MEASURES

Process measures explore the manner in which the health care provider interacts with the patient, they are a measure of physician skill. Information for this measure can be gathered through three general methods: actor-patients simulating a visit to a doctor's office, chart abstraction of medical records, and clinical vignettes (Peabody 2000).

Process measures are preferred as tools for assessing quality since they allow for the identification of acceptable and appropriate care of a patient. However, for quality of care measures based on process to be valid, there must be evidence that changes in the process result in improvements in patients' outcomes (Brook 1984, 1996, 2000, Siu 1992). Process measures also tend to be the most stringent on quality of care, since the omission of any of the required steps in the provision of care results in patients receiving less care than they should (Brook 2000).

There are a series of circumstances that make process measures more difficult to obtain in less developed countries. In addition to general deficiencies in information systems (De Geynt 1995, Durán-Arenas 1998), there is a specific dearth of medical records (Peabody 1994). Without the legal requirements for keeping accurate medical records, or the financial incentives for documenting procedures to link them to patient billing, it is unlikely that there will be a tradition of careful record-keeping in health facilities. In less developed countries, the lack of legal requirements and financial incentives for keeping medical records is exacerbated by the lack of information technology resources to integrate medical records from different sources, making it unfeasible to find accurate medical records that reflect the events of a visit to a doctor's office. This creates a difficulty in research involving the process dimension of quality in less developed countries.

OUTCOME MEASURES

The outcomes dimension of quality include a set of results from the delivery of health services (Batalden 1994, McGlynn 1995), most important among them:

- clinical status: blood pressure, lung capacity, hemoglobin value, etc.
- functional status: range of motion, ability to ambulate, cognitive skills, etc.
- quality of life: pain, energy, sleep, etc.
- satisfaction with care

The difficulty in using outcomes as measures of quality is that outcomes can depend on spurious relationships that don't reflect the level of care received by the patient (Brook 1996). They can also take too long to process and evaluate, for instance in conditions like breast cancer, where 5 year survival rates are a standard measure of outcome. By the time the results are compiled and processed, the institutions that provided the care may have changed markedly, making it impossible

to make any assessment about the quality of care they are providing (Brook 2000).

INTEGRATING THE DIMENSIONS OF QUALITY

Distinguishing between structural deficiencies and process shortcomings allows participants in the health care sector to chose the right set of strategies that can improve a health system. The process of care, for example, can be remedied through training, practice standards, licensing requirements, and completed in a time horizon of months to a few years. These interventions putatively cost less than some structural improvements, such as hospital construction or the implementation of a drug delivery system. On the other hand, buying a MRI unit takes less time and money than putting students through medical school or through a medical residency program. This relationship between structure and process, and how it works is important to identify.

Although conceptually the relationships between different attributes of quality of care and patient's health are well established, this dissertation aims to provide an empiric basis for exploring the magnitude of the impact of different levels of quality of care on health in a population. This is an integrative effort that recognizes the systems qualities of health care provision. The examination of one independent attribute at a time may not yield the same results as the integration of the different dimensions in a model, for it is reasonable to believe that additions in one of the dimensions of quality will have effects in the other dimensions as well. A new diagnostic method and device will be accompanied by training in the equipment.

The theoretical framework that will thus guide the research of this dissertation, is lead by a distinct view of the workings of process as "a set of activities *(that)* transform inputs (such as patients with health care needs, skilled health care staff, equipment, supplies, and

financial resources) into outputs (including information, diagnoses, and treatment decisions) that are associated with outcomes realized by patients and others who are affected by the care (families, employers, communities)" (Batalden 1994).

The systemic integration of the three dimensions of quality of health care begets a theoretical model whereby structure influences process and process influences outcomes, since it is through the process component that structural factors affect the health status of patients. This notional model will be tested at two different levels. The relationship between structure and process will be examined at the individual practitioner's level. The relationship between process and outcome will be examined at the population level. This will not examine the relationship between individual providers' practices and the health of their patients, but rather whether at the level of health facilities within a city, the application of a set of practice procedures is associated with better health outcomes of the community.

Figure 1.1 The Relationship between Structure, Process, and Outcome

Structure	Process	Outcome

Physician Characteristics :
- A g e
- E xperience
- T raining

Site:
- W orkload
- L ocation

Acess to Resources:
- O w n ership
- E quipment
- S upplies

Events in the Visit to the Health Care Provider

- Self-Assessed Health
- Clinical Indicators
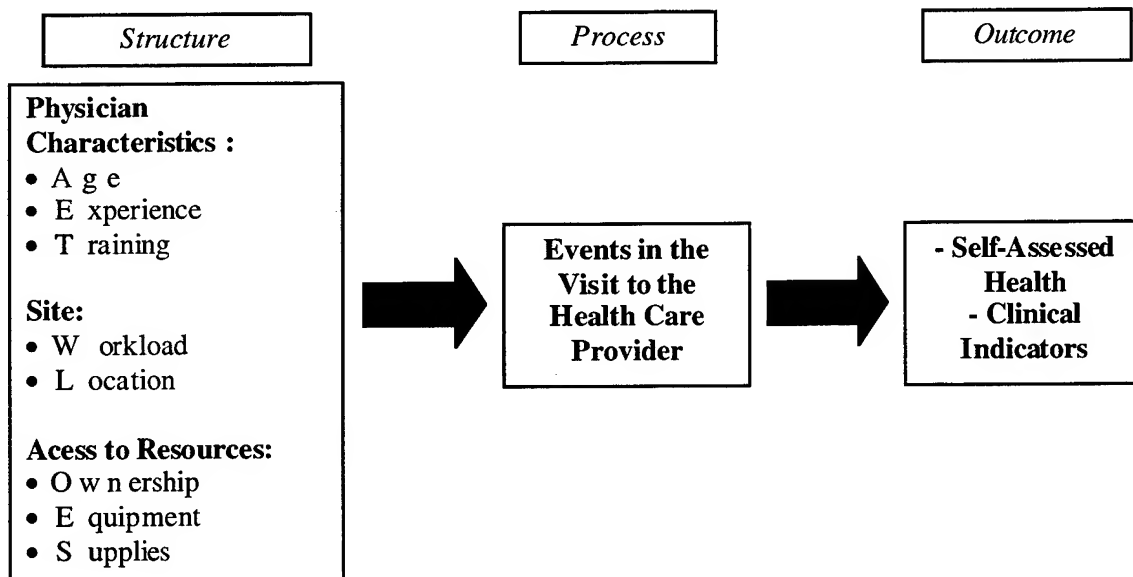
This dissertation will make assertions based on explorations of the changes in aggregate measures of process from a census of health care providers and relating this to aggregate measures of health from a geographically defined population.

The data to explore the specific linkage of individual providers and patients is not available to the author, thus this relationship will not be explored.

METHODS

STRUCTURAL MEASURES

The structural measures in this research originate with the surveys of health facilities and providers that were conducted in Macedonia. Structural Measures comprise physician characteristics and physical means-i.e. hospital, equipment-(Donabedian 1980). The instruments used in this study were designed to gather information on the number and condition of health inputs as well as characteristics of primary care providers. Specifically, it measures important clinic characteristics like: patient utilization rates; provider characteristics such as age, years of training, and specialty; structural measures such as capital equipment, supplies, and medications; staffing levels; costs of services, tests, and evaluations; and referral practices. The measures used in this study are a census of all the medical doctors, supplies and equipment present in the hospital, and contains all measures of structural quality (Donabedian 1980, Gertler 1990, Peabody 1995, Nordyke 2000).

All public physicians and a sample of private physicians who provide primary care in the selected municipalities are included in this component survey (Nordyke 1998). The Facility Survey measures important clinic characteristics: utilization rates, structural measures such as equipment and supplies including drugs, staffing levels, costs, and referral practices. It also reports on important provider characteristics such as specialization, gender, age, education, and years of experience. Thus, the Facility Survey provides detailed measures of the structural quality of primary care (Nordyke 1999). Further discussion of the survey design is provided in Chapter 3.

PROCESS MEASURES

As discussed previously, there are three general methods of gathering information on process of care: Actor Patients, Medical Chart

Abstraction, and Clinical Vignettes. In the case of Macedonia, as in most other less developed countries, using medical charts is not an option since the information systems that exist in developed countries are not ubiquitous (Durán-Arenas 1998). Even in the presence of working information systems, medical charts can contain omissions that underestimate the quality of care received (Luck 2000). This problem is only intensified in environments with deficient records systems.

Actor Patients, although considered the "gold standard" of process measurement (Rethans 1987, Colliver 1993, Pieters 1994, Badger 1995, De Champlain 1997, Colliver 1995, Luck 2000, Glassman 2000) are rather expensive and logistically complicated when compared to clinical vignettes. Furthermore, clinical vignettes have been validated in clinical practice and provide consistent indications of medical knowledge as well as reliable evidence of actual physician behavior (Peabody 2000, Dresselhaus 2000).

Clinical Vignettes

The clinical vignettes in this study are paper cases given to physicians in the United States and Macedonia. Criteria for scoring are based on international standards of care using process attributes known to result in better health outcomes.

Vignettes or written case scenarios have been used widely to measure process in a variety of practice settings and in different research areas including education, demography, and health services research (Sriram 1990, O'Neill 1995, Glassman 1997). Clinical vignettes have been used successfully in previous studies to assess quality of medical care (Shawyer 1996, Kalf 1996). They are easily administered, less costly than other methods and can be adapted to different types of clinical practices (Glassman 1996).

Although there is concern about the validity of written simulations (Jones 1990, Thaver 1998), the clinical vignettes used in this study

have been shown to provide consistent indications of medical knowledge and reliable evidence of actual physician behavior (Peabody 2000, Dresselhaus 2000).

The criteria found in the clinical vignettes rely on evidence-based medicine, which stresses the importance of evidence from clinical trials and does not rely on unsystematic clinical anecdotes or intuition (EBMWG 1992). All the criteria found in the clinical vignettes are based on processes that can be said to lead to better outcomes based on empirical evidence (i.e., evidence based medicine).

The clinical vignette scores, as a measure of physician skill, can be examined on the basis of physician characteristics or on the basis of health outcomes (McGlynn 1997), this study does both. On the basis of physician characteristics, this dissertation examines what affects the scores on the clinical vignettes. This part of the analysis explores the relationship between structure and process. The structural elements that affect process are identified based on the results to the clinical vignettes and the data from the facilities and health care providers. On the basis of outcome, the relationship between changes in health status and scores in clinical vignettes is tested empirically.

The vignette presents a physician-respondent scenario of a patient who has a specific outpatient clinical conditions seen by doctors in primary care. The evaluation measures consist of questions about how the doctor would conduct the consultation: taking a history, conducting a physical examination, ordering tests, diagnosing, and providing treatment plans. Criteria for scoring are based on international standards of care using process attributes known to result in changes in outcome.

The clinical vignette is presented in sections corresponding to areas that recreate the sequence of a visit to a doctor's office. Physicians are prompted for open ended responses in each section. Each section begins with the presentation of new information gained from answers to questions in the previous section. After answering one

section, the physician cannot return to the previous question to revise the answer, thus the new information cannot be used to change (and improve) their answer[1].

The physicians are presented with a problem scenario based on a typical patient visit and asked questions about the history they would take. After their answers have been written down, the results of the history are revealed and doctors are asked about elements of the physical examination that they would perform. They are then given the findings of the physical examination and asked about the radiology or laboratory tests they would order. Those results are also provided, after which they are prompted for a diagnosis and the treatment they would recommend. Once the physician has answered a question and moved on, it is not possible to return to a previous section to revise an answer.

Completed vignettes are scored by an expert nurse abstractor or physician who records the items that the physician completes successfully. Explicit scoring criteria for each of the different cases were developed based on international standards of care for prevalent conditions that are observed by primary care physicians. The candidate criteria was then submitted to expert panels of academic and community doctors who were both generalists and specialist for the conditions. Based on their recommendations and group consensus, the master criteria list was modified and finalized (Nordyke 2000, Peabody 2000).

Clinical Vignettes have been developed and validated for nine common health conditions that are best treated in a primary care setting: coronary artery disease (CAD), hypertension (BP), chronic obstructive pulmonary disease (COPD), diabetes (DM), prenatal care (Prena), tuberculosis (TB), contraception (Contr), and low back pain (LBP) for adult patients, and diarrhea (Diarr) and cough with fever (Cough) for children (Nordyke 2000, Peabody 2000).

---

[1] The author has administered vignettes to groups of doctors in different locations in the Los Angeles, CA area.

Two specific case scenarios are developed for each condition corresponding to a simple and a complex presentation of the disease. The complex case scenario has comorbidities that are not present in the simple case scenario.

The case scenario for simple coronary artery disease corresponds to a clinic visit by a 65-year-old new patient, who comes to the doctor's office for a follow up on a myocardial infarction he had 3 month before. A summary of the items in the scoring criteria for this scenario are presented in Table 1.1.

Table 1.1

CAD1 Scoring Categories

Domain	Category	Description
Hx	Thrombolytics	ER treatment
	Procedures	Invasive procedures
		Noninvasive proc's
	CHF	CHF in hospital
		CHF now
	Angina & Other symptoms	Chest pain
		Syncope or palp'ns
	Selected Risk Factors/ Comorbidities	DM or Htn?
		High cholesterol
	Prevention	Past preventive care
		Alcohol
	Drug treatment	Current meds
		Drug allergies
	Risk Factors	Family history
		Smoking history
		Diet
		Exercise
Px	Rule out CHF	JVD
		PMI
		Edema
	COR	Chest auscultation
	Lungs	Back auscultation
	Evaluate for PVD	Carotid bruits
		Peripheral pulses
Test	Cholesterol	Cholesterol
	EKG	EKG
	ETT/Cardio	Treadmill
		Cardiology referral
	Chest X-ray	Chest X-ray
Dx	MI Diagnosis	Small inferior MI
Tx	Aspirin	Aspirin
	Beta blocker	Beta blocker
	Prevention – Counseling	Stop smoking
	Follow up	F/U Appointment
	Prevention – Care	Preventive care

A set of clinical vignettes for different conditions and their respective scoring criteria can be found in Appendix A.

One obvious question is how to understand the vignette scores done in Macedonia with scores and criteria for other countries. For calibration purposes, this dissertation will also include the

presentation of results of clinical vignettes administered in a US clinical setting. The data from clinical vignette administration in two Veteran's Administration Hospitals will be presented as a gauge of the performance of clinical vignettes.

SURVEY DESIGN AND SAMPLES

The data used in this research originates in three surveys conducted in the Former Yugoslav Republic of Macedonia in 1997-1998. The three surveys are:

- A survey of facilities and physicians' characteristics, including clinical vignettes administered to physicians.
- An exit survey of patients visiting health care facilities to collect outcome and access measures.
- A household survey that was a follow-up to the exit survey and included members of the patients household.

Figure 1.2 presents the linkage between the surveys and their sample sizes. Further discussion on the survey design is provided in Chapter 3.

Figure 1.2. The Macedonian Health Surveys

Figure 1.2. The Macedonian Health Surveys
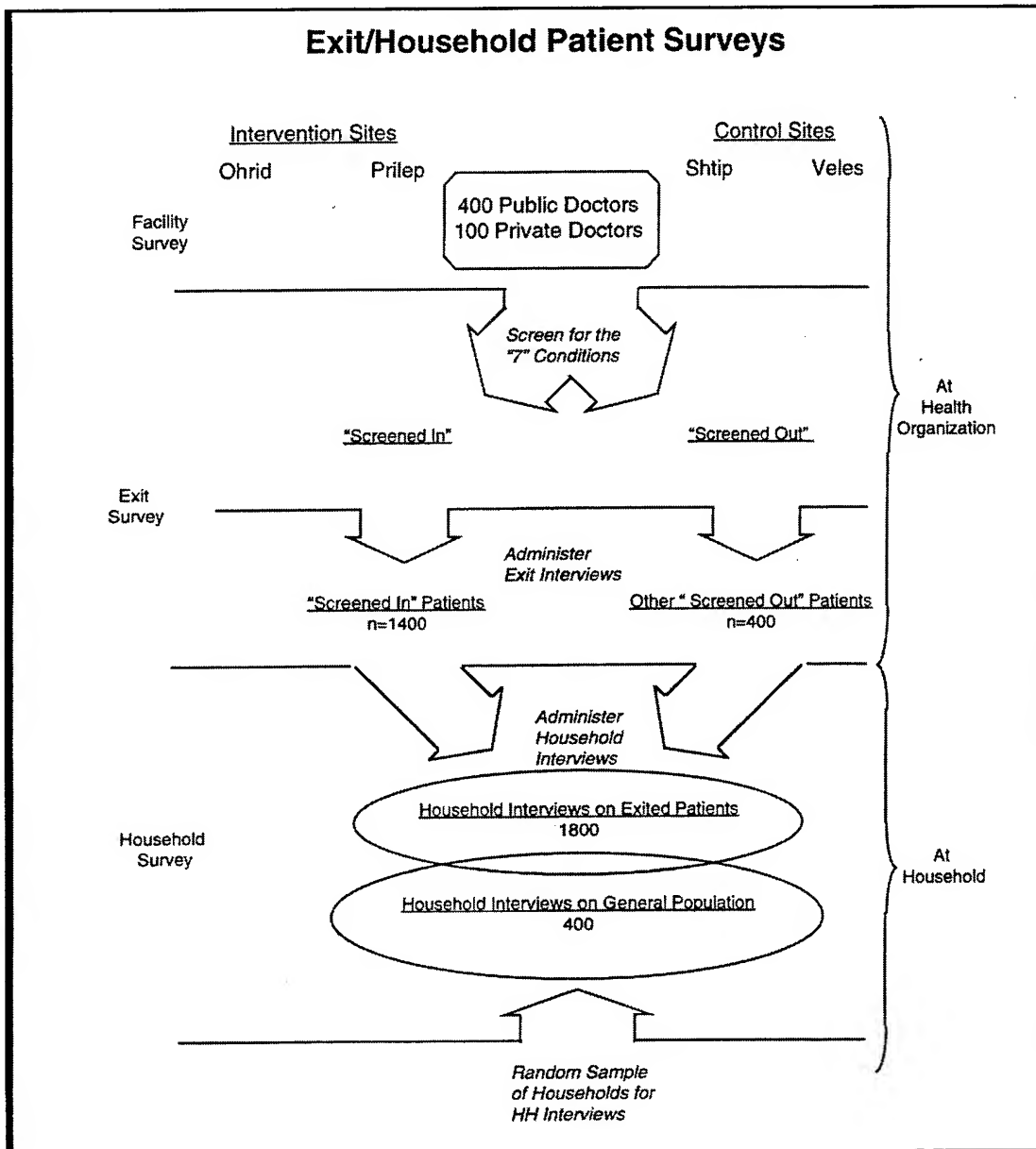
Source: Peabody 1999b

OUTCOME MEASURES

The outcome measures used in this study come from two Macedonian sources: the exit survey conducted at health facilities, and the household survey. The two measures used for analysis are the self-assessed health status and a set of clinical indicators of health outcomes.

The self-assessed health status contains the answer to the question: "In general, would you say your health is:" The five possible answers are: Excellent, Very Good, Good, Fair, Poor. This is used as a subjective indicator of outcomes, or health level.

The clinical outcome indicators used in this study are blood pressure, lung capacity, and blood glucose content. These three indicators are used to construct a measure of objective health or outcomes.

STATISTICAL TESTING

All statistical testing was conducted using STATA 7.0 statistical software and the SAS/STAT 6.12 software. All univariate statistical tests were two-sided, and were assessed for significance at the 0.05 level. For multivariate analysis, the corresponding p value is presented.

PART 1 DESCRIPTIVE ANALYSIS

The unit of analysis for all analyses in this section was the clinical vignette. The variation in vignette score by condition, type of condition, age of physician, specialty, and domain are examined. The clinical vignette scores are broken down into its domain components only in the analysis by domain of care. All other analyses use the total vignette score, aggregated across the different categories in the dimension of analysis.

Variations in vignette score by each measure were assessed by means of an omnibus one-way analysis of variance (ANOVA), with the measure of interest being the categorical class variable for each of the ANOVAs. If this test was statistically significant for a given measure, a series of follow-up two-sample t-tests was performed, with each t-test comparing the mean score for the segment of interest to the mean score for the remaining segments. Additionally, two sample t-tests were performed for all the different segments taking two at a time.

Analysis of variance (ANOVA) models are a statistical tool used to investigate the relationship between an observed variable and one or more dimensions in which it can vary (explanatory or predictor variables). ANOVA models do not require that any assumptions be made about the nature of the theoretical or statistical relationship between the observed variable and the explanatory variables, therefore there is no need to specify a function. The explanatory variables do not need to be quantitative, the observed variable however has to be quantitative.

PART 2 MULTIVARIATE ANALYSIS

All multivariate analyses will be performed using robust (Huber-White) estimators to account for clustering of vignette scores in facilities.

a)

Does Structure affect Clinical Process? The following model will be tested:

Process = f(Clinic Ownership, Physician Characteristics, Physician Workload, Location, Physical Infrastructure)

b)

Does Process affect Population Health Outcomes? The following models will be tested:

- Subjective: *Health Level = f (Process, Age, Gender, Education, Living Conditions, Access to Information, Distance to Hospital)*
- Objective: *Health Outcomes = f (Process, Age, Gender)*

SUMMARY AND ORGANIZATION OF THIS DISSERTATION

This study prospectively identifies variability in quality of care in a less developed country using data from the Former Yugoslav Republic of Macedonia[2]. Clinical vignettes are used to measure the process of care for ten common outpatient conditions: low back pain, diabetes mellitus, chronic obstructive pulmonary disease, and coronary artery disease, blood pressure, children's cough, contraceptive therapy, diarrhea, prenatal care, and adult cough.

After exploring quality variation, this research evaluates the factors that may affect the process of care in a less developed country. In this analysis we used the clinical vignette scores for the ten conditions specified previously together with data from a survey of health facilities in Macedonia to examine the relationship of structural aspects of care to clinical processes.

A final analysis looks at the variation of health outcome measures (subjective and objective) and how these relate to process and structural measures. This section will investigate how variations in process affect outcomes.

The culminating section of this dissertation provides a discussion of the observed results and recommendations for further study.

---

[2] Former Yugoslav Republic of Macedonia, FYRM, and Macedonia will be used interchangeably throughout the rest of this dissertation

TABLE OF CONTENTS FOR CHAPTER 2

QUALITY OF CARE IN DEVELOPING COUNTRIES

As pointed out by McGlynn (1997), different stakeholders in health systems have distinct definitions of quality, and thus differing expectations of the health care system. In developed countries, the quality of care literature is nourished by views from the variety of stakeholders, including patients, providers, and payers (McGlynn 1997, Shuster 1998), who have been engaged in the debate over quality for roughly two decades. In less developed countries, where the discussions are just beginning, the "payers" include government and providers too, and most important international lending agencies like the World Bank and the regional developing banks, who have significantly higher levels of resources than any stakeholder indigenous to the less developed country.

The literature on quality of care in developing countries that is available is limited and comes primarily from reports prepared by international agencies, or from studies financed primarily through international agencies' grants. This section will begin by considering very broad measures of quality, which although not necessarily investigated within the more focused framework of the dissertation, are helpful in providing an overview of the quality of health care in less developed countries.

In the developed world, the interest in measuring quality of care is driven by cost and an increasing awareness of the variability of care, and the realization that there are measures that can be taken to reduce variability and improve the quality of medical care (Eddy 1984, Weiner 1995).

At this general level these two factors are not different in the developing world. However, beyond the problems of variability in medical care, patients appear to face much lower levels of quality of care compared to developed nations (Mathiyazhagan 1998, Gani 1996,

Omaswa 1997, Pécoul 1999, Seidman 1998, Shawyer 1996, Thaver 1998). The literature on quality of health care in less developed countries identifies three general areas that are lacking:

- Access to Health Care Services
- Cost Efficiency
- Quality Orientation of Clinical Practice

This discussion will first reference factors not incorporated in the dissertation, like education and nutrition; second it will look at opportunity costs-it may be better to spend on education versus health; third it will consider the heterogeneity of inputs within regions; and fourth, it will parley the problem of measures-they are not the same in different regions, especially for inputs, but we will look primarily at process and outcomes.

The confluence of the three elements that are lacking in the quality of health care in less developed countries creates a level of care that is generally lower and likely produces lower health status in populations (Peabody 1999a). The evidence to support the claim that these three areas are lacking can be found in a few sparse empirical studies, as well as in development performance indicators provided by international agencies like the World Bank, World Health Organization, and the United Nations. Other areas that contribute to lower standards of care in less developed countries, such as female literacy, that although may be of great importance, are outside the scope of this work that focuses on measuring clinical care and improving care through policy interventions. Most important among the areas not explored here are: education, nutrition, poverty, and other community level variables such as leadership and infrastructure.

Another important point to keep in mind when comparing quality across different systems is that opportunity costs are different in societies with different levels of resources. Since opportunity cost is the highest valued alternative foregone in the pursuit of an activity--

the provision of health care in this case-it is reasonable to think of less developed countries as having other alternatives, like education,- that have not been met in an appropriate manner. Dedicating resources to providing health care competes with the provision of certain services that in developed countries are more completely addressed. Certainly the cost inefficiencies observed in developing countries may well also be present in countries with developed economies, however, the burden of those cost inefficiencies will have a greater marginal impact in countries with less resources.

Less developed countries have less homogeneity in some of the factors that affect health status, like nutrition, education, and other socioeconomic indicators and lower levels of resources that contribute to health--GDP and Medical facilities (Akin 1985, Klitgaard 1985). An exploration of the relationship between structure, process, and outcomes will elucidate the policy levers available to decisionmakers trying to improve a health system. Although conceptually all the studies support the idea that improvements in different areas of quality of care will improve health status, the contribution of the present study to the existing literature is to provide an empirical base for this relationship. The scarcity of factors compels us to explore the influences of structural and process quality indicators on the ultimate goal of a health policy intervention: health outcomes. Following the research framework of this dissertation, the relationships that will be tested, will be link between structure and process, and the link between process and outcomes.

One of the main problems in measuring quality are the metrics and the appropriate measures. For instance, when comparing costs, we cannot simply compare the cost of treating a cancer patient in country A versus the cost of treating a cancer patient in country B, especially when comparing developing countries with developed countries. In the poorer country the cost may be lower simply because the patient might receive less medical attention, or the cost of some of the inputs, like labor, may not be comparable.

Access to Health Care Services

Access to health care services in less developed countries differs from the levels observed in developed countries. As an illustration of the problems existing in poorer regions we can observe the causes of deaths in different regions of the World. Adequate health care services contributes to a reduction of deaths caused by preventable conditions.

Table 2.1

Age Adjusted Deaths as Percentage of Population

Region	Africa	The Americas	Eastern Mediterranean	Europe	South-East Asia	Western Pacific
Total Deaths	1.69	0.70	0.87	1.04	0.95	0.74
Communicable Disease, maternal and perinatal conditions and nutritional deficiencies	1.19	0.11	0.31	0.06	0.37	0.09
Non-communicable conditions	0.37	0.52	0.48	0.89	0.49	0.57
Injuries	0.13	0.07	0.08	0.09	0.09	0.08
Maternal Conditions	0.04	0.00	0.01	0.00	0.01	0.00

Source: World Health Organization 2001

Table 2.1 presents age adjusted deaths for different regions as a percentage of their total population. In Africa, deaths represent a higher percentage of the population than in any other region. The only other region where deaths represent more than one percent of the total population is Europe, although this is due to the high amount of deaths by non-communicable conditions brought by its aging population. A poor region like Africa, has a higher number of deaths as a proportion of its population caused by preventable causes such as communicable conditions, injuries, and maternal conditions.

To evaluate the access to medical care in developing countries versus developed regions, we can use the standard method of comparing structural indicators of quality such as availability of health facilities and services.

Table 2.2

Health Services and Use Comparison

		Low Income Countries	Middle Income Countries	High Income Countries
Inpatient admission rate % of population 1990-98		13	6	15
Average Length of Stay days		11	12	14
Outpatient Visits per Capita 1990-98		4	4	8
Child Immunization Rate	Measles	64	88	89
% of children <12 mo, 1995-99	DPT	70	88	91

Source: World Bank, 2001 World Development Indicators

Table 2.2 presents a comparison of health services and use for low, middle, and high income countries. By different measures, access to health care in low income countries is inferior to access to care in countries of the other two income categories.

The difference in immunization rate of children is another indicator of the lack of availability of health services in poorer countries. This lack of immunization is translated into a higher proportion of childhood diseases (WHO 2000).

The population in less developed countries allot a higher portion of their income to receiving health (Gillis 1992, Peabody 1999a), but many times the existing health services are of such poor condition, that patients are willing to pay even more in order to assure adequate care.

In Ghana, a survey conducted in 1998 showed that among children under five years of age who had presented diarrhea symptoms in the two weeks preceding the survey, only 27.2 percent received advice or

treatment from qualified health care workers, while 31.2 percent received advice from pharmacists, druggists or other sources, and the rest (41.6 percent) received no advice or treatment (GSS 1999); this is an example of poor quality of medical care

An illustration of how much patients are willing to part with in hopes of receiving better quality of care is an example of introduction of private health services in Kenya, where most private clients of the largest managed care organization are willing to pay for health insurance levels that cost US$205 to US$262 a year, despite the fact that average per capita income in the country is US$340.00 (IFC 1999).

Access to medical care includes access to appropriate medications. This is a problem as well in less developed countries. (Pécoul 1999) identified more than 10 conditions prevalent in developing countries—including tuberculosis, diarrhea, and malaria—where the appropriate drugs either have a price that is prohibitive for patients in poor countries, are only available on a limited basis, or are simply not produced because there is no commercial interest in producing them. This finding corresponds to the conclusion of other studies of drug availability in other less developed countries (Foster 1993).

As international lending agencies such as the World Bank and the International Monetary Fund progressively tie their lending programs to restructuring public spending and to eliminating excess spending, health services are often cut creating new barriers to access to health care. International lending agencies may not make additional loans to the country unless privatization and cutbacks of public sector services are made (Stoker 1999, Wouters 1992, PAHO 1995, WB 1993). If the political leadership is not able to rechannel diminishing public resources to care for the vulnerable, health outcomes may suffer (Peabody 1999a). Just like in the United States the concerns about managed care have focused on the barriers of access to vulnerable groups and cutbacks in spending for clinical services, the implementation of "austerity plans" in

developing countries may be placing more people at risk of not receiving adequate health care (Stoker 1999).

In Chile, almost a quarter of the patients covered by the government-subsidized managed care organization ISAPRE, choose to receive services in lower quality public clinics and hospitals every year because they cannot afford the necessary copayments. In 1997 in Argentina, public hospitals reported that approximately 1.25 million outpatient visits by patients who had medical coverage through privately administered funds. Before seeking the hospitals, patients had encountered problems making their copayments, or bureaucratic confusions in coordinating a provider (Stoker 1999).

Cost Efficiency

When comparing the quality of costs, it must be asked: cost efficiency, compared to what? Unlike precise mechanical systems, efficiency measurement in health services tends to be a measurement of how much improvement is possible, rather than a comparison to an agreed standard. In the case of cost efficiency, the question of interest is whether the allocation of inputs is ideal in term of desired outputs.

The problem of cost efficiency has two aspects: allocative efficiency and technical efficiency. The former refers to the spending that is distributed compared with other uses for public funds (often outside of the health sector), the latter concerns itself with the way that different interventions or health services are administered-perhaps comparing one treatment option to another. Although they are separate aspects of cost efficiency, they are nonetheless closely related.

Allocative Efficiency

Table 2.4

Health Expenditure Comparison

		Low Income Countries	Middle Income Countries	High Income Countries
Health Expenditure as a % of GDP	Public	1.2	2.5	6.0
	Private	3.1	2.6	3.7
	Total	4.5	5.0	9.7
Health Expenditure Per Capita PPP $		74	267	2,587

Source: World Bank, 2001 World Development Indicators

Allocative efficiency can be illustrated through the figures in health expenditure provided in Table 2.4. The figures in the table show that low income countries spend significantly less on the health of their population, as a percentage of GDP and on real terms. The share of GDP destined for health in low income countries, is less than half of the share dedicated to that purpose in high income countries. Most of that share—more than 70%--is borne by private sources in low income countries. By contrast, in high income countries public resources finance more than 60% of health expenditures. These low expenditures are associated with higher child mortality, deficits in children's growth, and other negative outcomes (Brentlinger 1999, Pebley and Goldman 1999).

Technical Efficiency

Technical efficiency examines the way in which different interventions are administered; one area of interest is in the efficiency in the use of wage and non-wage inputs and technologies.

Table 2.6

Burden of Disease by Cause and Demographic Region, 2000

	WORLD	INDIA	CHINA	OTHER ASIA AND PACIFIC ISLAND COUNTRIES	LATIN AMERICA AND THE CARIBBEAN	MIDDLE EASTERN CRESCENT	SUB-SAHARAN AFRICA	FORMERLY SOCIALIST ECONOMIES OF EUROPE	ESTABLISHED MARKET ECONOMIES
DALYS (%s)									
I. Communicable, maternal, perinatal, and nutritional conditions	35.6	44.2	12.5	32.9	27.4	39.4	60.6	6.3	7.2
A. Infectious and parasitic diseases	20.5	25.9	3.8	17.4	15.1	16.7	41.2	1.9	3.7
1. Tuberculosis	2.9	6.8	1.0	2.2	1.1	1.2	4.3	0.4	0.1
2. STDs excluding HIV	1.0	1.5	0.0	1.7	0.9	0.4	1.7	0.5	0.3
3. HIV	3.0	3.3	0.1	2.3	4.2	0.2	6.4	0.2	2.7
4. Diarrhoeal diseases	5.5	6.7	0.8	5.2	3.7	8.2	9.4	0.2	0.2
diseases	4.0	4.1	0.5	3.0	2.3	4.7	8.9	0.0	0.0
6. Bacterial meningitis and meningococcaemia	0.3	0.3	0.3	0.3	0.3	0.4	0.3	0.2	0.1
7. Hepatitis B and hepatitis C	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0
8. Malaria	2.0	0.2	0.0	0.8	0.3	0.2	7.8	-	0.0
9. Tropical cluster diseases and leprosy	0.5	0.6	0.0	0.2	0.5	0.1	1.2	-	0.0
10. Leprosy	0.0	0.1	0.0	0.0	0.0	0.0	0.0	-	-
11. Dengue	0.0	0.1	0.0	0.1	-	-	0.0	-	-
12. Japanese encephalitis	0.0	0.0	0.1	0.1	-	-	-	-	-
13. Trachoma	0.1	0.0	0.1	0.0	-	0.1	0.1	-	-
14. Intestinal nematode infections	0.2	0.2	0.4	0.6	0.4	0.1	0.1	0.0	-
B. Respiratory infections	6.3	8.0	3.1	6.1	3.4	9.0	9.1	1.5	1.2
C. Maternal conditions	1.1	1.3	0.4	1.3	1.0	1.4	1.8	0.5	0.2
D. Perinatal conditions	5.1	6.1	2.7	5.0	5.5	8.6	5.9	1.7	1.4
E. Nutritional deficiencies	2.6	2.9	2.5	3.1	2.5	3.8	2.6	0.7	0.6
II. Noncommunicable	47.4	38.7	69.1	50.5	54.6	44.9	21.0	75.7	81.8
A. Malignant neoplasms	6.6	3.8	12.7	7.1	5.6	3.1	2.5	13.0	15.9
B. Other neoplasms	-	-	-	-	-	-	-	-	-
C. Diabetes mellitus	0.8	0.8	0.5	0.8	1.5	1.0	0.2	0.9	2.3
D. Endocrine disorders	0.4	0.0	0.3	0.3	1.3	0.7	0.4	0.3	1.1
E. Neuro-psychiatric conditions	12.2	9.6	15.6	14.3	18.3	10.6	5.0	16.4	24.5
F. Sense organ diseases	1.0	1.5	1.3	1.3	0.8	0.8	0.8	0.1	0.1
G. Cardiovascular diseases	11.4	11.4	13.6	12.0	9.6	12.9	4.2	25.5	19.0
H. Respiratory disease	5.3	3.7	13.4	3.1	4.5	4.6	2.8	6.2	5.1
I. Digestive diseases	3.4	2.3	4.2	5.1	3.9	4.0	1.7	4.3	4.7
J. Genito-urinary diseases	1.1	0.8	1.2	1.1	1.2	2.0	0.8	1.4	1.1
K. Skin disesases	-	-	-	-	-	-	-	-	-
L. Musculo-skeletal diseases	1.6	0.7	2.1	1.6	3.8	0.7	0.4	4.4	4.5
M. Congential anomalies	2.5	3.4	3.1	2.4	2.4	3.1	1.6	1.7	1.7
N. Oral conditions	0.7	0.5	0.6	1.0	1.2	1.2	0.2	0.8	0.9
III. Injuries	17.0	17.2	18.4	16.6	18.0	15.6	18.3	17.9	11.0
A. Unintentional injuries	11.9	15.0	13.3	13.6	12.7	7.9	10.7	12.3	7.9
B. Intentional injuries	5.1	2.1	5.1	3.0	5.3	7.7	7.6	5.6	3.1

Source: World Bank, 2001 World Development Indicators

The study of technical efficiency fits within the research framework of this dissertation because we want to explore how inputs (structure) are processed to produce outcomes. The difference in technical efficiency in developing countries versus developed countries can be illustrated with the figures from Table 2.6. The percent of disease burden that is accounted for by communicable diseases is an indicator of the stages of development of a particular region. Leaving aside China and the Formerly Socialist Economies of Europe, where their economic systems and demographic policies have created unique conditions, we see that less developed economies have a higher disease burden associated with communicable diseases than developed economies.

There are two main factors outside of health that influence the distribution of disease burden. One is the population demographics. In societies with a significant fertility decline, it is more likely that we will find a heavy disease burden associated with the conditions that affect adults. The contribution to burden of disease of conditions like cardiovascular disease and cancer becomes more important in adulthood and increases with age—in some cases exponentially (SOV 1999). On the other hand, societies that have fertility in excess of replacement, will have a heavy component of their burden of disease associated with conditions that are more likely to affect children: communicable diseases. The demographic composition of society is thus an important driver of the burden of disease distribution through the different epidemiological profiles associated with different stages in life (Evans 1994, Peabody 1999a).

The other factor affecting the distribution of disease burden is precisely the technical efficiency of controlling conditions that can be either prevented or whose incidence can be significantly reduced with relatively low cost interventions such as better education, improved hygiene and an adequate household environment, unlike non-communicable diseases such as cancer or heart disease. The fact that the prevalence of conditions whose prevention can be achieved through lower cost

2-11

interventions is relatively higher, is also an indication of the inefficiencies in allocative policy of health resources.

Quality Orientation

The data on structural measures can only provide us with part of the picture: the availability of resources and the organization of care. It does not provide us with information on the quality of those resources. We cannot for example ascertain anything about the conditions of the visit to the clinic, or how well trained the doctors are, or whether they provide the correct treatment. Moreover, in all developing countries, physicians and hospitals tend to be concentrated in particular urban areas-a problem of structural access to health care-so these health indicators provide even less information on the health services available to the entire population.

There is empirical support that the people in less developed countries demand higher quality care. In a study that contained a survey of 1,000 households in rural India, Mathiyazhagan (1998) found that physical accessibility to health care services was a major factor in willingness to pay for a private health care provider and a willingness to join a rural health insurance scheme. The survey results also suggest that the existing government health care provider's services are not easily accessible.

In low income countries, the mere presence of access to health providers and other structural inputs in the area is not necessarily a promise for appropriate health provision. In a study of childhood malnutrition in 27 villages in El Salvador (Brentlinger et. al. 1999), it was found that even though the majority of the villages (81%) had qualified health workers, none of the existing programs of the health workers provided food when child malnutrition was detected. In another survey of public and private health clinics in Jamaica, Peabody (1994) found that even though private clinics tended to be in better condition, better equipped and supplied, and better able to provide appropriate

laboratory tests and results in a timely manner, they still provided inferior prenatal diagnosis and counseling, and less family planning services than public clinics.

Another study looked at the clinical skills of foreign medical graduates and found evidence of common skill deficiencies in graduates of foreign medical schools (Ziv et. al. 1998). Related studies have shown that even when there are no serious knowledge deficiencies in doctors in developing countries, their prescribing practices are not close to accepted international medical standards. A study of 201 private practitioners in the slums of Karachi, showed that the doctors could diagnose 3 out of 4 conditions presented to them through clinical vignettes (Thaver 1998). However, at least 1 in 10 practitioners prescribed a non-accepted treatment; almost two thirds of them agreed that they were over-medicating their patients (Thaver 1998).

These gaps in clinical practices and thus the quality of care may be attributable to idiosyncratic beliefs of the patients, lack of knowledge in prescribing practices, and changes in prescribing practices that have occurred since the time when the doctors were trained (Thaver 1998, Pebley et. al. 1996, Shawyer et. al. 1996).

Deficiencies in regulatory mechanisms (Bhat 1996) and, when they exist, a lack of awareness of regulatory principles (Bhat 1999), are problems all too common in developing countries that also directly affect the quality of care provided by health practitioners. One reason may be the absence of tradition of continuing medical education programs that are tied to registration or continuation of practicing licenses (Bhat 1999). This can promote obsolescence of medical knowledge and thus a general lower quality of care.

Ultimately health care services are valued according to their marginal contribution to health status. If there is a low contribution to health status after a visit to a health provider, then it is not

likely that there will be higher utilization of health services, even when economic factors are not barriers to access (Wouters 1992).

The literature on health economics considers that health is produced from a series of inputs, among them the services furnished by the health care provider (Arrow 1963, Grossman 1972, Akin 1985, DaVanzo 1990, Newhouse 1992, Robalino 2001). The willingness to pay for such services will depend on their marginal productivity and expected return. As patients develop a belief about how a visit to a health care provider will contribute to an improvement in their health, their probability of making the visit changes (Mwabu 1986). For example, in a multiple visit study in Kenya, Mwabu (1986) found that if patients come to believe that their illness cannot be removed through the treatment received in a modern medical clinic, they are less likely to visit a medical health care provider and more likely to visit a traditional healer.

It thus becomes necessary to gain a better understanding of the workings of process in less developed countries, since empirical studies have shown that in those regions the demand for medical services is highly insensitive to price (Akin 1981, 1986, Bitran 1989, Dor 1987).

The Contribution of This Study

Assessing variation in quality of care provided in the context of a health system and identifying the sources of variation is a difficult task. Previous work, at the population level as has been presented above, looks at access, efficiency, and limited assessments of quality. These different perspectives provide an approximate picture of the health status of the population and the staggering disparities that exist between what is done and what could be done.

What the current literature on quality of health differentials is not presenting, is the critical link between the different structural attributes of quality mentioned (access, cost efficiency), and the quality of clinical practice. Making matters worse, we do not have information from less developed countries on the relationship between the quality of clinical practice and the ultimate goal of any health intervention: better health outcomes. Although conceptually the studies support the existence of a relationship, there is no substantial empirical evidence for this.

This study will present an empirical exploration of the relationship between structural aspects of quality and clinical practice; and between clinical practice and health outcomes. This is a unique contribution to the understanding of the workings of quality that can allow for the formulation of policy recommendations to improve health status in developing countries. Do structural aspects influence clinical practice? Does clinical practice matter? These questions will be addressed in the analysis of data on the clinical practice of a group of doctors and the health outcomes of their patients.

TABLE OF CONTENTS FOR CHAPTER 3

I. QUALITY MEASUREMENT STRATEGIES

Quality of medical care has three different dimensions:

- Structural Measures—physician characteristics and physical means (i.e. hospital, equipment, access to health care).
- Process Measures—the factors of the encounter between the health care providers and the patients.
- Outcome Measures—the subsequent health status of the patient or population.

This study relies on data for all three different dimensions of quality of care to present the analysis. The data was obtained in three surveys conducted in the Former Yugoslav Republic of Macedonia in 1997-1998. The three surveys are:

- A survey of facilities and physicians' characteristics, including clinical vignettes administered to physicians.
- An exit survey of patients visiting health care facilities to collect outcome and access measures.
- A household survey that was a follow-up to the exit survey and included members of the patients' households.

The three surveys served the purpose of collecting different information that indicates the level of quality of care at the level of the community. The survey of facilities and physicians provided information on the physicians and their practice environment. The exit survey provided information on health status and on access to medical care. The household survey provided health status on users and non-users and looked at the other important factors that determine health (e.g. education and income).

The clinical vignettes were administered as part of the physicians' survey. The clinical vignette scores, as a measure of

physician skill, can in turn be related to structural measures—physician and practice characteristics-to explain vignette scores—or on the basis of health outcomes—vignette scores can be used to explain outcomes. This study does both in order to empirically explore the framework developed in Chapter 1 of the relationship between structure, process, and outcome.

The survey of facilities and physicians provides data that allows us to analyze the structural measures (physician characteristics and physical infrastructure) to identify the relationship between material inputs and determine how they might affect process as measured by the clinical vignettes.

The vignettes used in this study provide data on clinical practice. Vignettes were given to doctors for conditions that are both prevalent in Macedonia (and the United States) and have a significant burden of disease. These diseases or conditions are best treated in primary care settings.

The study used vignettes for the following conditions: coronary artery disease, hypertension, chronic obstructive pulmonary disease, diabetes mellitus, prenatal care, tuberculosis, contraception, and low back pain for adult patients, and diarrhea and cough with fever for children. The vignettes for all conditions were used in Macedonia, the US study only collected information on four conditions: coronary artery disease, chronic obstructive pulmonary disease, diabetes mellitus, and low back pain for adult patients.

The evidence on the prevalence of these diseases in Macedonia comes from primary data collection. Data on prevalence were collected in Macedonia using a nonrandom exit survey preformed 12 months prior to the study (Nordyke 1999). These results were analyzed with estimates from the United States. Based on these results, Nordyke (1999) estimates that these conditions contribute or cause over half of all deaths in adults in Macedonia, a proportion not too dissimilar from developed countries such as the United States; in addition, they account for a

significant burden of disease in Macedonia. The data for the two countries is summarized in Table 3.1.

Table 3.1

Summary on Data for Five Chronic Conditions in the United States and Macedonia

Condition		Prevalence Among Elderly in U.S. (%)	Estimated Prevalence Among Elderly in Macedonia (%)
Hypertension	HTN	60	57
Coronary Artery Disease	CAD	16% M; 11% F	15
Chronic Obstructive Pulmonary Disease	COPD	11.0%	18
Diabetes Mellitus	DM	10%	13
Hypercholesterolemia	HC	28% M; 43% F	N/A

Source: Nordyke 1999

The US data is used primarily as a calibration of disease prevalence in Macedonia. Similarly, to calibrate clinical practice we use vignette data for the United States. The US vignettes have the significant advantage of having been also calibrated against actual clinical visits (Peabody 2000). The purpose of the US study was a comparison of clinical vignettes with other methods of measuring process: actor patients, and clinical chart abstraction. The clinical vignettes proved to closely match clinical practice and are a reliable and valid method of assessing process of care (Peabody 2000).

The culminating analysis in this study relates the variation of clinical practice to differences in outcome measures (subjective and objective). The information for the outcome measures is provided by the exit survey (subjective and objective) and the household survey (subjective and objective). The exit survey is a choice based sample that allows us to establish a link between the care of a specific condition within a clinic and the health outcomes of the patients that visit that clinic for that particular condition. For a thorough discussion on the distributional properties of choice-based samples, see Gill (1988), and Vardi (1985).

The patients that were surveyed receive their medical care in the facility, and thus establishing a connection between the vignette scores of the physicians in that facility for a particular condition and the outcome measures of the patients for those conditions, will allow us to determine how process affects health outcomes.

This study also provides a broader generalization of the effects of process on health outcomes for the populations. This is done through a survey strategy that collected health outcomes measured in the household portion of the survey design. The household survey also contained data from people who had not visited the clinic. There was also a random selection of households that had not gotten care. An association between the clinical vignette scores of the physicians from a particular facility, with the health outcomes scores from households served by that facility will present evidence of a link between the process of care and health levels in a community.

The survey strategy used in this study is illustrated in Figure 3.1 and shows the different components of the survey and their interaction.

Figure 3.1   The Macedonian Health Surveys
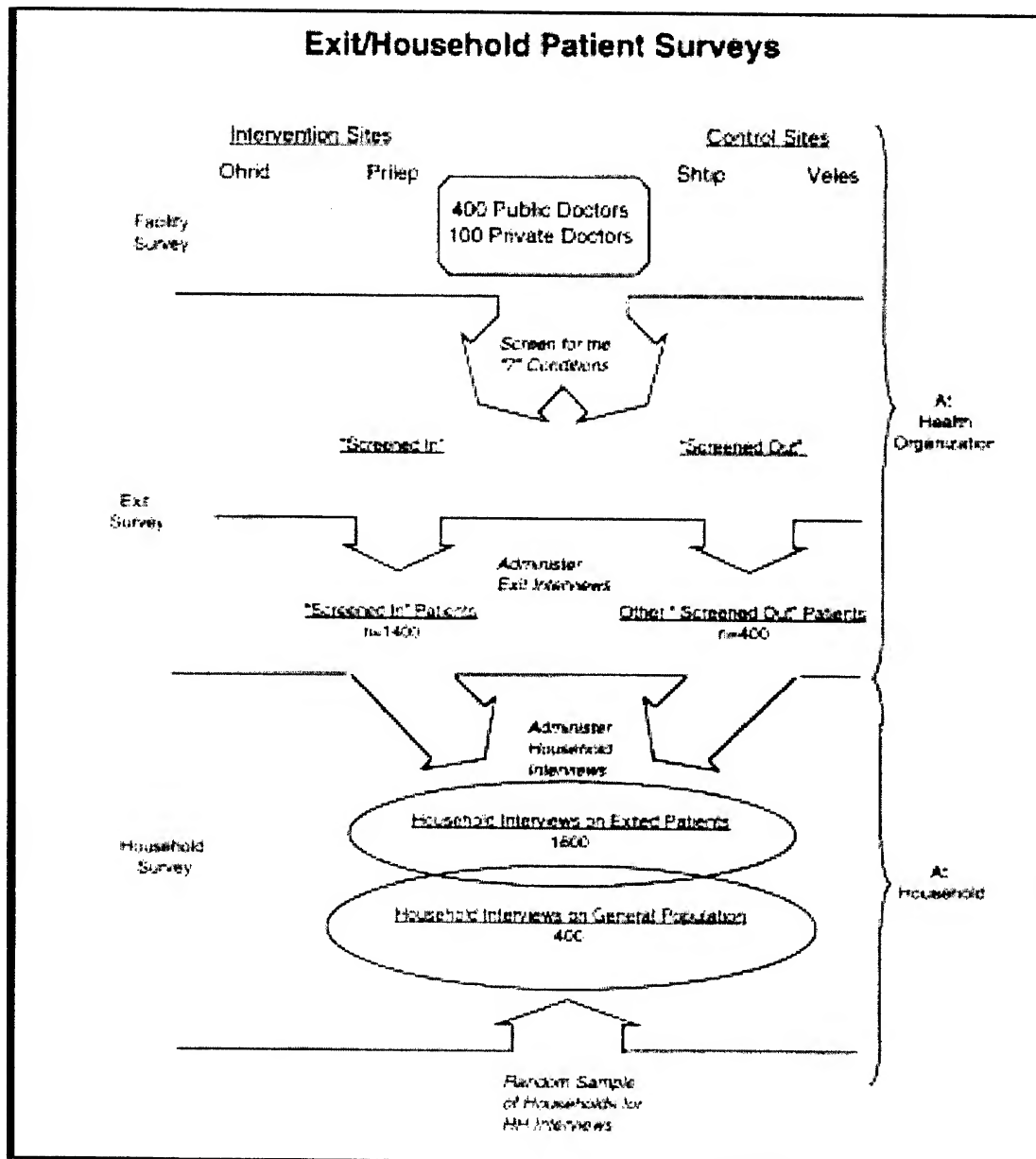
One of the main assumptions that permits us to make this inference is that we are dealing with a population that receives its medical care locally. Only 3.3% of the population surveyed traveled outside of their municipality to receive care (Nordyke 2000). This permits us to associate the level of quality of care in a geographic area to the observed health outcomes. This research design allows us to explore relationships at the community/population level, which is the area of interest of this dissertation. We cannot make assumptions at the physician/patient level, since the data collection mechanisms don't allow for the connection of individual providers with their patients. An analysis at the physician/patient level however, is not the intention sought at the policy level, where the interest is on aggregate measures.

II. THE SITE

Former Yugoslav Republic of Macedonia

The Former Yugoslav Republic of Macedonia (FYRM) declared independence from a disintegrating Yugoslavia in 1991. Although by 1992 it was established under its own constitution as a sovereign state, international recognition was delayed by Greece's objection to the new state's use of what it considered Hellenic name and symbols. It was only after the compromise name of FYR Macedonia was accepted that Greece lifted its trade blockade and agreed to normalize relations in 1995. After that the FYRM enjoyed full international recognition as a sovereign state. Macedonian was proclaimed a language in 1944, although Greece doesn't recognize this name. The language uses the Cyrillic script and is closely related to Bulgarian.

With a total area of 25,333 sq. km, the FYRM is home to 2,041,467 people (July 2000 est.) where ethnic Macedonians represent the majority (67%) but other groups are represented as well—Albanians (23%), Turkish (4%), Roma (2%), Serb (2%) and others (2%). These groupings also corresponds to the language segmentation of the country and is manifested in the religious distribution—67% Macedonian Orthodox, 30% Muslim, and 3% other religions. The ratio between the sexes is one to one.

With an unemployment rate of 35% (est. 1999), the economy can still meet the basic alimentary needs of the population, but the country completely depends on outside sources for its oil and gas, as well as most of its modern machinery and spare parts. Real GDP growth rate was estimated at 2.5% in 1999, largely held down by the severe regional disturbance caused by the Kosovo conflict. A map of the republic of Macedonia is shown in figure 3.2.

Figure 3.2 Map of the Former Yugoslav Republic of Macedonia

Source: CIA Worldfact Book 2001

Countries neighboring FYR Macedonia are: Albania, Serbia with Kosovo & Montenegro, Bulgaria and Greece. The capital of the Former Yugoslav Republic of Macedonia, Skopje is in the north of the country, on the Vardar River. Skopje, then called Scupi, was the capital of the kingdom of ancient Dardania. The Serbs and then the Turks came through in the late Middle Ages, and it became an important city in the Ottoman Empire. In 1963, an earthquake destroyed most of the city, including many ancient mosques.

Health Sector Reform

The Former Yugoslav Republic of Macedonia initiated reforms in its health care sector soon after its independence in 1991 as part of World Bank supported reforms throughout their economy. The health sector reform is comprehensive. It is generally directed at streamlining provision and improving services. Specifically to this aim, the reform included legislative initiatives, the purchase of new equipment, and the acquisition of management information systems. It also provided for the provision of continuing medical education that will be based on the baseline assessment of quality. Before independence, the health care system was part of that of the Socialist Federal Republic of Yugoslavia (SFRY).

The most fundamental parts of the health reform could be undertaken after the summer of 1996, when The World Bank awarded the Ministry of Health of Macedonia a loan of US $19.4 million—the Health Sector Transition Project Credit. The main goals of the reform is to provide universal access to high quality primary health care for the country's population, and the establishment of cost-effective financial and delivery systems.

Prior to 1991, the health care system was similar to that of other socialist countries, with a national planning system that oversaw the whole the health sector, down to the level of physicians' practices. The private provision of health care has been legal since 1991, although administrative barriers to entry have kept most of the health care provision in public hands.

The main purpose of the Macedonia's health sector reform is to improve the efficiency and utilization of primary health care. To do this, two policies have been proposed:

1. Increase consumer incentives for primary care by:
- setting coinsurance rates for specialty care higher than that for primary care;

- reducing copayment exemptions for all but the poor and the elderly; and

- creating a new Basic Benefits Package. This package would reduce overall expenditures and improve or maintain health status by shifting resources and patients to more cost-effective outpatient care.

2. Introduce capitation payment to primary care providers. This scheme, which marks a sharp departure from the previous system of fixed salaries in the public sector, provides significant incentives to doctors to improve their use of resources and the quality of care they provide.

Health Status

Among the republics and regions of the SFYR, Macedonia ranked second to last behind only Kosovo in GNP/capita, infant mortality, and life expectancy, in the 1980s. Like other post-socialist countries in Central and Eastern Europe, Macedonia suffers from a host of problems in its healthcare system, including: poor health status of the population, over-specialized and under-employed doctors, inefficient and undercapitalized clinics, and nontransparent financing systems that are not conducive to sound management.

Currently Macedonian health indices are improving, it's infant mortality rate having fallen to almost one fourth of its 1980's level to 13.4 deaths per 1,000 live births. This is lower than all of its bordering countries but Greece, where the infant mortality rate is less than one half of that of Macedonia—6.5 deaths/1,000 live births.

Life expectancy, at 74 years is comparable to higher income countries in Europe. But so are the leading causes of death—cardiovascular disease and cancer. However, the gains in population health status that Macedonia has achieved have been expensive. As illustrated in Figure 3.3, Macedonian public spending on health in terms of GDP is high relative to both western and eastern European countries.

Figure 3.3 Comparison of Public Health Expenditures
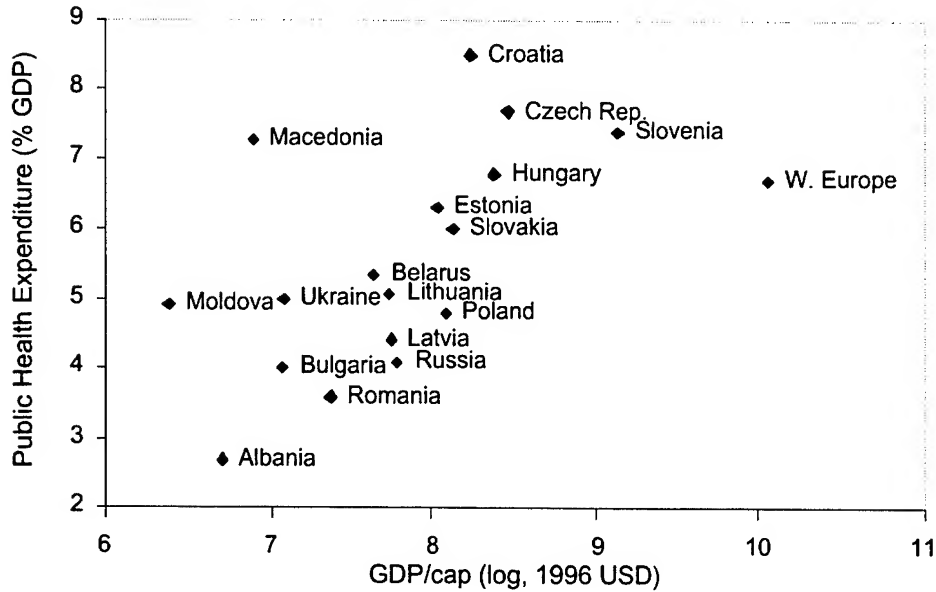
Source: Nordyke 2000

Even though Macedonia has one of the lowest GDP per capita in its region, the share of GDP destined to health care is substantially higher than most of the countries in the area. Among its neighboring countries, only Greece spends a larger share of its GDP health, but in Greece it is mainly driven by private expenditure. While in Macedonia 85% of the expenditure on health as a percentage of GDP is borne by the public, in Greece public outlays only account for 57% of the expenditure on health as a percentage of GDP.

In Macedonia, as in other less developed countries, average health status indices are not necessarily reliable indicators given the disparities in health status between regions of the country. Tables 3.2 and 3.3 provide an illustration of the variation in health indicators across the country.

Table 3.2

Regional Disparities in Health Indicators

	IMR	Live Births	Infant Deaths	% Urban
Republic of Macedonia	22.5	33,487	752	58.7
Debar	32	663	21	54.5
Shtip	31	704	22	82.4
Skopje	28	8,793	245	81.4
Kriva Palanka	27	301	8	28.8
Tetovo	27	4,067	108	44.5
Kavadartsi	11	568	6	0
Vinica	10	293	3	52.4
Berovo	9	234	2	49.5
Kratovo	7	140	1	59.8
Probishtip	5	204	1	3.8
Sveti Nikole	4	250	1	62.1
Demir Hisar	–	115	–	0

Source: Peabody 1997

Table 3.3

Ethnic Disparities in Health Indicators

Ethnicity/Age	0-4	5-14	15-24	25-44	45-64	65+	Total
Macedonians	354	33	55	138	890	6262	894
Albanians	771	36	38	106	978	5572	558
Turkish	499	25	50	117	967	5503	576
Rhomas	644	20	46	189	2322	8333	748
Vlachs	0	0	0	41	129	2774	484
Serbs	199	45	0	155	966	4878	902
Other	3403	49	137	326	2167	11856	1668
All Groups	571	33	50	135	951	6184	814

Age-specific death rates by ethnicity, 1994 (deaths per 100,000)
Source: Peabody 1997

As Tables 3.2 and 3.3 suggest the variation in health indicators across the country is significant. This variation is not only geographically based, but it is also influenced by the different ethnicities in the country.

Table 3.4

Insurance Coverage by Expenditure Segments

Expenditure Quintile	Number of Respondents	Percent Covered
0-20	860	55.6%
20-40	858	57.9%
40-60	856	61.8%
60-80	860	66.5%
80-100	855	67.7%
Total	4,289	62.1%

Source: RAND Household Survey Supplement

Just as the variation in outcomes is observed in Macedonia for different geographic and ethnic segments of the population, economic divisions are present as well. Data from households in Macedonia support the notion that access to care is dissimilar for the different economic tiers in the country, as illustrated in Table 3.4. Even though these differences may not be currently as substantial as those observed in other less developed countries, as market reforms are implemented and economic disparities widen, access to health care may become even more disparate.

Even though the health sector in Macedonia has similarities with those of developed countries-highly trained physicians, low incidence of communicable conditions, high percentage of GDP devoted to health-there are other areas that are more similar to less developed countries. Particularly accounting for the fact that it is a country in transition, where some of the current highlights in the health sector have been inherited by a planned, albeit inefficient, economy. In summary, Macedonia suffers from a great deal of the problems plaguing the health sectors of less developed countries. Low availability of resources, poor health status of the population, heterogeneous access to health care, highly inconsistent health indicators among population segments, and inefficient management of health resources are some of the problems that contribute to an overall picture of a health system where improvements can be made.

Capitation Evaluation Program

Prior to nationwide implementation, the government sought to demonstrate and evaluate the health reforms. They were demonstrated in a pilot program in two large municipalities: Ohrid and Prilep. An additional purpose of this pilot program was to collect baseline data that will allow the measurement of the impact of the health reforms.

This pilot program, referred to as the Capitation Evaluation Program (CEP), was conducted by the International Project Unit (IPU) and RAND. The CEP provided an ideal opportunity to collect data on the different components that constitute measures of quality of care.

The demonstration sites, the municipalities of Ohrid and Prilep, were chosen by the Ministry of Health early in the planning phase of the reform program. Based on census data from the National Statistical Office (NSO) and a previous Household Survey conducted by RAND in October 1996, the two control sites were selected on demographic variables important to health status and utilization—Shtip and Veles. Table 3.4 provides a summary of the health assets in the selected municipalities.

Table 3.5

Health Facilities in the Selected Study Sites

| City | Intervention Sites | | Control Sites | |
	Ohrid	Prilep	Shtip	Veles
Number of Primary Care Doctors	79	105	40	54
Health Centers	1	1	2	1
Medical Centers	1	1	1	1
Other Urban Health Facilities	14	0	3	10
Rural Stations	5	10	9	8
Private Group Practices	7	3	5	17

Source: Nordyke 1999

Selected demographic variables for the intervention and control sites, as well as other municipalities in Macedonia are presented in Table 3.6.

Table 3.6

Demographic Variables for the Intervention and Control Sites

Municipality	Population (000s)	Ave. Quarterly HH Expenditures (MKD)	Fraction Urban	Ave. Age (yrs.)	Fraction w/ Insurance
Intervention sites					
Ohrid	60.8	40,726	.63	36.3	.65
Prilep	93.2	48,833	.75	36.4	.61
Control sites					
Shtip	50.5	51,668	.78	38.1	.67
Veles	65.5	46,043	.70	33.4	.60
Others					
Kavardarci	41.8	42,690	.67	35.1	.53
Kumanovo	126.5	46,885	.46	31.1	.57
Strumica	89.7	58,489	.78	38.1	.69

Source: Nordyke 1999

Based on the analysis of an extensive array of health system indicators for all municipalities in Macedonia, the four sites included in the pilot program were considered the most comparable (Peabody 1997).

III. THE FOUR DATA COLLECTION INSTRUMENTS

All data from Macedonia used in this research was collected through the efforts of the RAND Capitation Evaluation Program (CEP) team working under the Health Sector Transition Project credit for Macedonia, supported by the World Bank.

The data used in this dissertation corresponds to the Round 1 data collection effort—collection of baseline data—which was completed by December of 1998.

The field work for the surveys was managed by RAND/IPU staff and by staff from the National Statistical Office (NSO)) in Skopje, the capital of Macedonia. Before the fieldwork was started, the survey teams underwent training in Skopje. Survey training included general as well as survey-specific information. It contained training on sampling procedures, quality control procedures, as well as administrative procedures. Survey training was conducted using written materials, lectures, and presentations by survey management staff. Practice included planned exercises, role-playing, evaluation and testing, and field exercises with actual respondents.

The surveys-Facility, Exit, and Household—were conducted in four municipalities: Ohrid, Prilep, Shtip, and Veles. The clinical vignettes were administered as part of the Facility Survey. Figure 3.4 shows the three surveys and their relationship.

Figure 3.4   The Macedonian Health Surveys
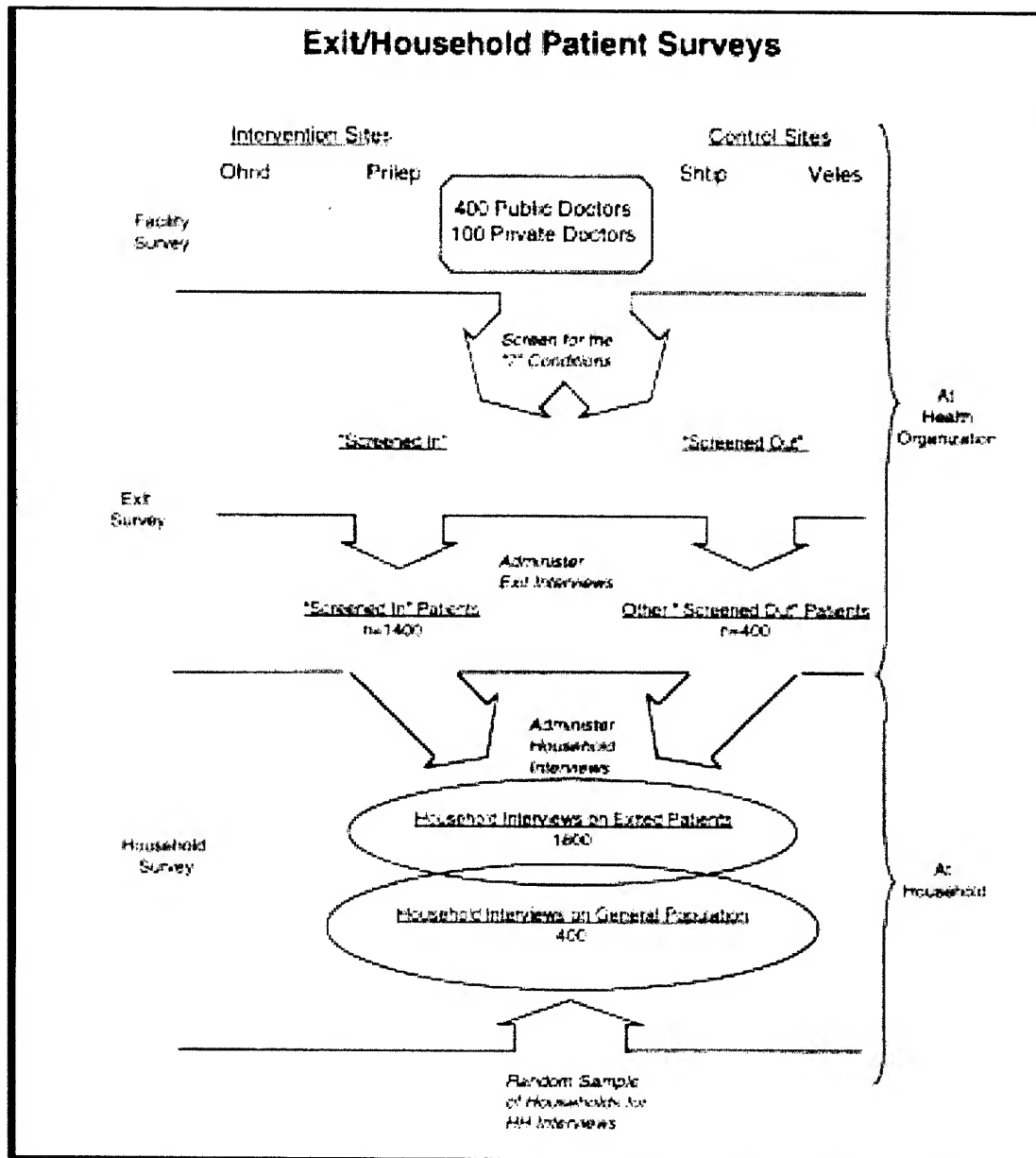
Source: Peabody 1999b

The CEP relied on a sophisticated research design (see Figure 3.4) and required the cooperation of several groups to assemble a team capable of conducting the field work and data management tasks. The CEP team pretested and revised the four survey instruments, performed the

data collection, and completed the data entry and cleaning. Figure 3.5 presents a schematic of the team structure.

Figure 3.5 Structure of CEP Research Team

Function				
Design and Training:	RAND/IPU Staff NSO			
Site Supervision/ Coordination:	Ohrid Research Supervisor NSO Supervisor	Prilep Research Supervisor NSO Supervisor	Shtip Research Supervisor NSO Supervisor	Veles Research Supervisor NSO Supervisor
Facility & Exit Data Collection:	9 Facility / Exit Interviewers	11 Facility / Exit Interviewers	6 Facility / Exit Interviewers	7 Facility / Exit Interviewers
Household Data Collection:	5 Household Interviewers	5 Household Interviewers	5 Household Interviewers	5 Household Interviewers

Source: Nordyke 1999

The RAND/IPU staff provided overall survey management and logistics support. The central NSO staff in Skopje also served this function for the Household Survey. Within each study municipality, research supervisors headed up each field work team and were responsible for the progress of the fieldwork teams in their municipality and the quality of field data collection. Each research supervisor managed a team of up to 10 interviewers, depending on the size of each municipality.

A separate team, managed by the NSO, conducted the Household Survey fieldwork. In each of the four study sites, the NSO municipality supervisor and the research supervisor formed a supervision/coordinator leadership tandem and worked closely together to ensure coordination of tasks and information transfer in the Exit and Household Surveys.

Facility and Physician Survey

Sample Frame

A census sample frame was developed by drawing on several sources of data on physicians. The medical centers in the four municipalities supplied their current employment registries, showing the names and locations of all public doctors practicing in their boundaries. This was checked against the registry that the Project Unit maintained in the capital, Skopje, to develop the master list for public doctors. The Ministry of Health and the Chamber of Physicians each provided registries of doctors licensed to practice privately across Macedonia. These two datasets were merged to build the Master List for private doctors in the study sites. Participation, although voluntary, was over 90 percent in the four municipalities. Of 274 eligible physicians in the four municipalities, 273 participated in the survey.

Table 3.7

Facility Survey Sample

Municipality	Public Health Centers and Polyclinics	Private Clinics	Public Rural Clinics	Other Public Urban Clinics
CEP Intervention Sites				
Ohrid	43/12	16/14	5/11	8/8
Prilep	52/10	8/7	11/10	14/5
CEP Control Sites				
Shtip	38/10	6/5	3/12	1/1
Veles	37/13	19/6	11/13	7/7

Source: Nordyke 1999
(# of Physicians/# of Facilities)
Note: The number of physicians sum to 279 due to double counting of some Physicians who work concurrently on health centers and rural clinics on a part-time bases

Implementation

The Facility and Physician Survey measures important clinic characteristics: utilization rates, structural measures such as equipment and supplies, staffing levels, costs, and referral practices. This survey also includes an observations section where the interviewer

provides his or her own information on the observed condition of the facilities. The facility and physician survey contains a census of all public primary care doctors and the facilities where they practice as well as a sample of private primary care providers. It therefore reports on important provider characteristics such as specialization, gender, age, education, and years of experience.

All private and public physicians who provided primary care in the CEP municipalities were included in this component survey. Thus, the Facility Survey provides excellent measures of the structural and process dimensions of the quality of primary care.

The Facility Survey was developed early in 1997. The first draft of the facility instrument was translated and back-translated in April and May 1997. Pre-testing of the Facility Survey in the Macedonian language began in clinics in Skopje (and hence by facilities that were not part of the study) in May 1997, and continued in outlying municipalities in June and July 1997. In total, over 30 doctors completed the Facility Survey in the pre-test phase. Based on the pre-test, several questions were reworded primarily due to translations of medical terms.

At the time Round 1 of the CEP went to the field, there were about 221 primary care physicians in the 4 survey municipalities. The organization of these doctors varied significantly between study sites. In Shtip, for example, there were 33 doctors in the urban Polyclinic and a total of 40 public doctors in the municipality. In contrast, the Health Center in Prilep housed 47 doctors while Prilep had a total of 78 doctors in public practice. The departmental structure varied significantly between study sites. Rather than sampling doctors based on practice or department parameters, the Facility Survey is designed as a census of primary care doctors in the survey municipalities. A total of 317 primary care doctors were surveyed in the baseline.

Survey researchers used the Master List to arrange office visits with all public and private doctors. All eligible doctors in the sample frame completed Round 1 of the Facility Survey, a 100 percent response

rate. Over 95 percent of interviews were completed during the initial appointments. The remaining interviews that were partially completed, were finished within two days of the original appointment. Following each interview, the interviewer reviewed and edited the questionnaire. As the questionnaires were completed, the forms were delivered to an Assistant Research Supervisor for a second level of editing. Any errors were promptly returned to the field for correction or re-interview if required. Prior to delivery to the National Statistical Office for data entry, the Research Supervisor edited 20 percent of all facility questionnaires, returning questionnaires to the field for correction or re-interview as required.

The Facility and Physician Survey contained the following sections:
Screening Questions
Section A: Physician
Section B: Equipment
Section C: Services and Activities
   Community-Based Health Promotion Activities
   Laboratory Services
   Equipment
   Supplies
   Department Medications
Section D: Family Planning Activities
Section E: Financial and Administrative Information
   Patient Mix
   Fees for office visits and lab tests
Section F: Clinic Staff
Section G: History of the Facility
Section I: Direct Observation (completed by the Interviewer)
   Utilization
   Condition of the General Examination Room
   Places for Vaccine Storage
   Laboratory
   Pharmacy Checklist

This constitutes an exhaustive list of all physical means in the health facilities.

Clinical Vignettes

Sample Frame

The Clinical Vignettes were administered as part of the Facility Survey. Therefore the same sample frame was used in selecting the physicians that completed vignettes.

Implementation

The main research tool for the measurement of the process of care used in this paper is the clinical vignette—a written case where the physician states the course of action to be taken based on a presented scenario.

One feature common to all the vignettes used in this study is a focus on clinical conditions that are not only currently prevalent in Macedonia—representing a significant burden of disease (Peabody 1997)—but also diseases or conditions generally considered to be best treated in primary care settings.

Following the Facility Survey, Clinical Vignettes were administered to all primary care doctors in the study municipalities. In each municipality, a group administration session was arranged with the director of the local health center. Doctors were invited to attend either a morning or afternoon session depending on their shift in the clinics. As expected, attendance by public doctors was quite good. In all sites, additional visits were required to allow private doctors to complete the vignettes. In two of the sites, Ohrid and Prilep, two separate days were required to allow all public doctors to participate in the vignettes.

All physicians participating in the Facility Survey completed vignettes in the following conditions: coronary artery disease, hypertension, chronic obstructive pulmonary disease, diabetes, prenatal care, tuberculosis, contraception, and low back pain for adult patients, and diarrhea and cough with fever for children.

The clinical vignettes were translated into the Macedonian language, and back-translated in May 1997. Pre-testing of the Clinical Vignettes in the Macedonian language began in clinics in the capital, Skopje, in June and July 1997. The vignettes were not administered in any sites outside the capital to reduce the possibility of contamination of the sample. In total, over 40 doctors completed Clinical Vignettes in the pre-test phase. Based on the pre-test, several scenarios and corresponding questions were reworded due to awkward translations of medical terms.

Extensive pre-testing confirmed by preliminary results from the CEP shows that the revised vignettes worked well cross-culturally. It was also concluded that the level of information requested in the vignettes is appropriate and that Macedonian providers do not have problems using the vignettes or identifying critical information when they know the answer to the question. Focus groups showed that there was no difficulty understanding the case or what was required to successfully complete the clinical vignette.

Vignettes were administered and completed in Macedonian. Macedonian physicians, trained in research methods and experienced with the vignettes, translated individual doctors, answers to the vignettes. To ensure uniformity of translations, the four translators translated the same 4 vignettes in an effort to standardize terms. Additionally, each of the subsequent individual translations was reviewed by a colleague for quality control and consistency. When translations were complete, they were shipped to the US for scoring by a single medically trained record abstractor and data entry. Uncertainty over clinical diagnostics or therapeutic terminology (in an estimated 5 percent of responses) was resolved by a team of Macedonian and U.S. physicians.

Exit Survey

Sample Frame

Nearly 50 percent of the patients completing Exit Surveys also completed the same Household Survey. This was done to obtain an identical set of household covariates as for the sample from the general population. In total, 1800 households who had at least one member completing the Exit Survey also completed a Household Survey. Four hundred households were also randomly selected from the national census office in the four municipalities. By having household data from both the exiting patients (1800) and a random sample of all households (400), it was possible to identify for selection effects between users and non-users.

Implementation

Patients leaving the medical facility were asked to complete a questionnaire. During the interview, patients were screened for the seven adult medical conditions. The Exit Survey, as originally designed, included all 9 previously mentioned conditions. However, pre-testing showed that questions about contraception were too sensitive to yield reliable responses, so this condition was consequently dropped. Diarrhea was also dropped from the survey protocol because the incidence of children seeking care for diarrhea was very low, thereby leaving 7 conditions for tracking by means of the Exit Survey: heart disease, chronic obstructive pulmonary disease, hypertension, diabetes, tuberculosis, prenatal care, and cough with fever (for children under 5 years of age).

Those who "screen in" received an outcomes section containing questions about their health condition, identical to that provided in the Household Survey for the general population. In total, over 1800 patients who were "screened in" for at least one of the conditions completed the Exit Survey. In addition, 400 patients who "screened out" with none of the conditions were asked to complete the survey. This was

done to evaluate for any selection effects between those who used the facility for one of the nine conditions versus those who used the facility but did not have any of the conditions in the vignettes.

After providing the responses about their condition-specific outcomes, patients were asked whether they consent to three physiological tests. The tests were: blood sample, blood pressure, and lung capacity.

The response rates for the Exit Survey exceeded 85 percent. The administration of the physiological tests was conducted in a safe and voluntary manner. The tests were described in detail to the patients and a consent was obtained. The interviewers used standard, portable, medical test equipment designed for home use. Each of the interviewers had received prior training in the safe and correct usage of the testing equipment.

The Exit Survey was developed early in 1997. The first draft of the exit instrument was translated and back-translated in April and May 1997. Pre-testing of the Exit Survey in the Macedonian language began in clinics in the capital, Skopje, in June 1997, and continued in outlying municipalities in July 1997. In total, over 100 patients completed the Exit Survey in the pre-test phase. The pre-test validated the sampling scheme for selecting Household Survey participants from the Exit Survey respondents.

The calculation of the required sample size for the Exit Surveys to detect significant effects is illustrated using the example of hypertension. It is estimated the mean systolic blood pressure in hypertensive patients to be 165 mmHg with a standard deviation of 17 mmHg. It is further assumed that an effective treatment will yield a 10 mmHg drop in systolic BP for an individual; a successful treatment is one that achieves this improvement in at least 50 percent of the population. Thus, in a population of hypertensives, we are trying to detect a 5 mmHg drop in systolic BP; an effect size ($z = D / s$) of 0.30. Assuming that the sample size is not negligibly small and that the effect size is not large (does not approach 1.0), then the required

sample size is 360 under a two-tailed distribution.   Table 3.8
summarizes these calculations for seven conditions.

Table 3.8

Summary Calculation for Chronic Condition Sample Sizes

Condition	Measure and Mean Value	Std. Dev. (s)	Clinically Sig. Change (indiv.)	Clinically Sig. Change for Population (D)	Effect size (z)	Min sample size
Hypertension	Systolic BP: 165 mmHG	17mmHg	10 mmHg	5 mmHg	.30	360
Coronary Artery Disease	Annual number of exacerb.: 1.8 (log mean)	1 (log sd)	50%	25%	.45	160
Chronic Obstructive Pulmonary Disease	Lung Flow: Men: 600 l/min Women: 460 l/min	M: 48 W: 42	10%	5%	M: .63 W: .55	106
Diabetes	Blood sugar: 225	70	35	17.5	.25	500
Fever and Cough (Adult)	Fever and cough: 90% of resp.	N/A	N/A	10%	.38	220
Fever and Cough (Child)	Fever and cough: 90% of resp.	N/A	N/A	10%	.32	310
Prenatal Care	Birthweight : 3232 g	555	N/A	150	.27	430

Source: Nordyke 1999

The sample sizes listed in the last column represent the total
number of patients who have the corresponding condition; the required
sample sizes in each municipality are one-quarter of the total.   Based
on these parameters, the minimum sample sizes for each location were
computed.   Table 3.9 presents the minimum sample size required and the
actual number sampled. As shown in Figure 3.3, a minimum of 400 patients
in the Exit Survey report that they have none of the conditions.

Table 3.9

Sample Sizes by Condition and Municipality

Condition	Minimum Sample Size Goal/ Municipality	Municipality			
		Ohrid	Prilep	Shtip	Veles
Hypertension	90	129	147	168	107
Coronary Artery Disease	40	108	105	116	83
Chronic Obstructive Pulmonary Disease	30	65	62	53	47
Diabetes	125	103	141	125	99
Fever and Cough (Adult)	55	36	58	33	58
Fever and Cough (Child)	55	63	88	63	60
Prenatal Care	110	57	113	111	110

Source: Nordyke 1999

Questionnaires were drawn at random from the different groups. Names, addresses, and other identifying information from these questionnaires were delivered to the Statistical Office for administration of the Household survey. Up to 10 questionnaires were drawn in this way, depending on the number of questionnaires in each group. Following a draw, the questionnaires with multiple condition responses were redistributed into another condition category so that they might be selected based on another condition. In this way, a total of 50 patients per disease were selected in all four municipalities, yielding 1,400 patients with at least one of the conditions. Another 100 patients without any of the conditions were selected in a similar way in each municipality for a total of 400 more. Through this sampling process, 1,800 patients were selected for administration of the Household Survey.

Household Survey

Sample Frame

The Household Survey and the Exit Survey are closely linked; the bulk of the Household Surveys in the CEP are administered to patients completing the Exit Survey. The selection of households from the Exit Survey respondents was described above. Here, we describe the sample design for the general population.

Selection is based on the National Statistical Office standard practice for household surveys, which is a two-stage stratification. The nation is divided into geographical units, "enumeration districts," which contain an average of 70 households each. Enumeration districts (EDs) are classified on two dimensions: capital/non-capital and urban/rural. EDs with fewer than 10 households in rural areas and less than 20 areas are excluded from the sample. Once stratified, households are then randomly sampled within the eligible enumeration districts to attain the desired sample size. For example, for the quarterly Household Survey conducted by the National Statistical Office, 5 households are sampled in each ED, yielding 1,050 households.

For the CEP, rural households were oversampled at a relative sampling rate of 1.3-to-1 in the general population household survey across all 4 municipalities. In other words, a given rural household had a 30 percent greater chance of being selected than a given urban household in the same municipality. This allows the detection of urban versus rural effects that are 5 percent smaller than those that would otherwise be detectable without oversampling, at the cost of requiring that other effects be only 0.7 percent larger than the previously minimum detectable size.

Implementation

The household health surveys collected information on detailed household social, economic, and other demographic characteristics as well as health information for the individual household members. In

addition to the outcomes section administered to the patients in the exit survey, the household survey collects information on patient satisfaction, self-assessed health status, and insurance information. All members of the household were interviewed. After their information was obtained, patients were asked about their utilization of health care services and their health status. This was using standard health survey questions of Form SF-36, which have been widely validated for construct validity (Hays 1990, Keller 1998, Litwin 1998, Wasson 1999, Page 2001). A unique feature of this survey is that objective health status measures were obtained. Blood pressure was measured, random glucose measured and peak pulmonary flow rates assessed.

There were two types of surveyed households: those randomly selected from the general population and those with a member who completed the Exit Survey of primary care patients. Household members in the general population sample receive the same physiological tests as those administered to respondents in the Exit Survey. Households from the Exit Survey sample did not receive the tests.

There were two types of households included in the survey:
- Households with a member that completed the Exit Survey of primary care patients
- Households randomly selected from the general population

The randomly selected group served as a control and allowed to account for any selection bias that might exist (or not exist) had only the Exit Survey group been examined.

Household members in the general population sample were administered the same physiological tests as completed by the respondents of the Exit Survey. Participation was strictly voluntary, and so is the participation and responses of individual household members.

In the general population sample, in addition to gathering standard individual and household covariates, the interviewers used a

sophisticated "health outcomes section" that screens respondents for self-reported heart disease, chronic obstructive pulmonary disease, hypertension, diabetes, tuberculosis, pre-natal care, and cough with fever (for children under 5 years of age).

These health outcomes measure the way citizens with at least one of 8 medical conditions are treated and their access to care.[1] For example, respondents with diabetes are asked about medications, frequency and results of glucose tests, whether a doctor has advised them about diet or exercise, and the cost of their care.

One of the advantages of a survey design using a random sample of the general population is that it is possible to estimate the level of undiagnosed hypertension, diabetes, and lung disease by using the checks of blood pressure, blood glucose, and lung capacity. In Round 1 of the CEP, 100 randomly selected households in each of the four study municipalities completed the Household Survey for the general population. Given an average of about 4.5 persons per household, approximately 1,800 individuals are included in this sample.

---

[1] Questions about diarrhea were retained in the outcomes section in both the Exit and Household Surveys, although diarrhea patients were not tracked. Thus, health outcomes are measured for only seven conditions.

US DATA

The US data used in this paper was collected in a study carried out between December 1996 and August 1997 in the primary care outpatient clinics of two large urban Veterans' Administration medical centers. Only vignette data was collected in the US sites.

Sample Frame

All the residents in their second and third year at the medical centers as well as the attending faculty were eligible for the study (n=98). There were only three refusals. Vignettes were administered to 40 randomly selected providers (20 at each site) for four conditions, for a total of 160 clinical vignettes scores.

Clinical Vignettes

The vignettes used in the US were the scenarios corresponding to 4 common outpatient conditions: low back pain (LBP), type II diabetes mellitus (DM), chronic obstructive pulmonary disease (COPD), and coronary artery disease (CAD). Each condition then had a scenario reading presenting a complex version and one presenting an uncomplicated version, for a total of eight vignettes. The complex versions included a comorbidity, required additional diagnostic evaluation, or indicated more involved treatment.

The US Clinical Vignettes are used as an illustrative calibration of how the instruments perform in a clinical practice setting in the US. The purpose of showing this data is to explain by example how the validated instruments perform and to provide a gauge to consider the results in Macedonia.

TABLE OF CONTENTS FOR CHAPTER 4

ANALYSIS OF CLINICAL VIGNETTE SCORES AND SELECTED VARIABLES

A.1. QUALITY OF CARE IN MACEDONIA—VIGNETTE RESULTS.

Clinical Vignette Scores. As presented in Chapter 1, clinical vignettes were evaluated based on the physicians' responses and how correctly they addressed the clinical practice items deemed necessary for the respective condition presented. Each physician in the study was scheduled to complete four clinical vignettes representing one different condition each. Therefore, each physician produced up to four different vignette scores, which are used as individual observations for the purposes of analysis. This analysis takes the percentage correct items scored on each of the vignettes completed to construct a score that can potentially vary from zero to one.

In this analysis, only non-zero total scores were evaluated, since complete scores of zero were considered non-responses. However, it was possible to receive a score of zero in a particular domain of care, provided that the physician had answered the rest of the vignette. The analysis thus was based on 809 valid clinical vignette scores representing 207 physicians. The analysis of vignette scores against selected demographic variables is not aggregated by physician. Table 4.1 presents a distribution of the valid clinical vignette scores.

Table 4.1

Distribution of Valid Clinical Vignette Scores

City	Nr. Of Completed Vignettes	Nr. Of Doctors With Valid Vignette Scores	Nr. of Doctors by Nr. of Valid Vignettes Completed		
			4	3	<3
Ohrid	210	54	50	3	1
Prilep	184	47	43	4	
Shtip	161	41	38	3	
Veles	254	65	60	4	1
Column Total	809	207	191	14	2

Vignette Scores in Macedonia. The mean vignette score in Macedonia was 48.0%, the lowest score was 5.8% and the highest score was 84%. The 95th percentile of the scores laid below 83.3%. 400 clinical vignettes (49%) had a score below 50%, and 79 clinical vignettes (10%) had a score below 30%. Figure 3.1 illustrates the distribution of the vignette scores.

Figure 4.1 Distribution of Clinical Vignette Scores

Distribution of Vignette Scores
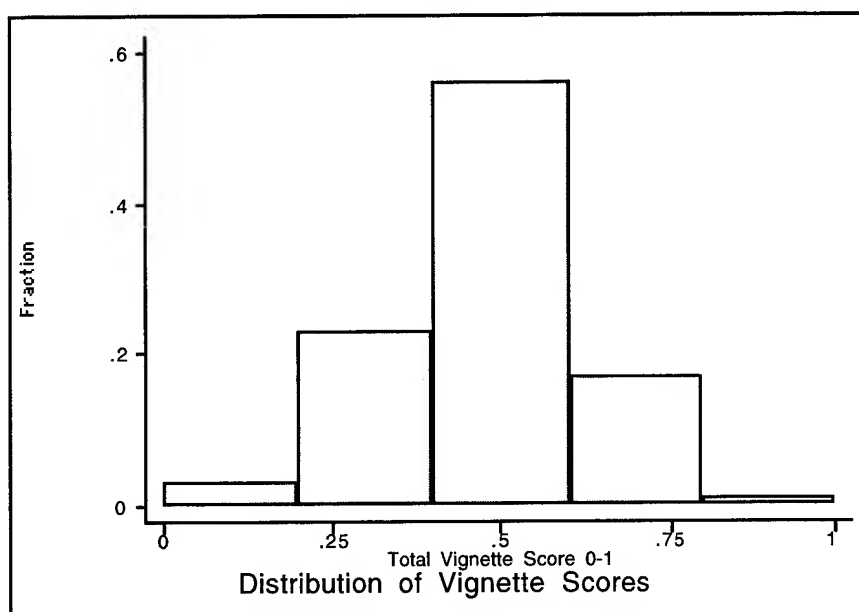
Figure 4.2 Clinical Vignette Scores by Condition
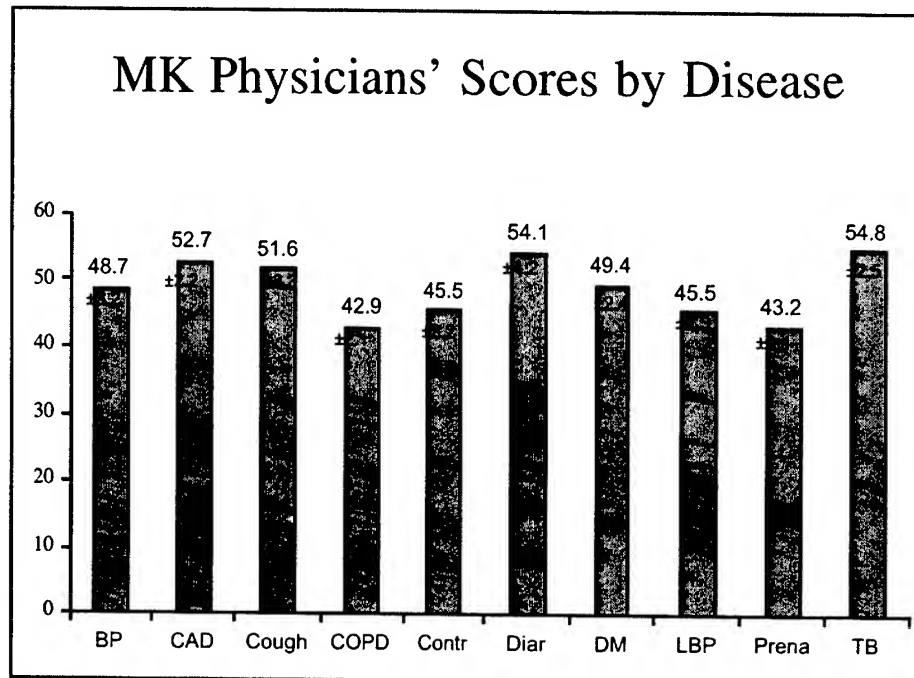
Variation by Condition. Clinical vignette scores are presented as percentage of items correct, and a 95% confidence interval for the score is provided. The overall vignette scores for physicians in Macedonia varied significantly by condition (p<0.001), as illustrated in Figure 4.2. The highest mean vignette score was for TB (54.8%) followed by Diarrhea (54.1%). The lowest mean score was for COPD (42.9%).

When compared to the mean scores of the remaining conditions, doctors scored significantly higher in 3 conditions—TB, CAD, Diarrhea. Doctors scored significantly lower in 2 conditions—COPD, Prenatal Care—when compared to the mean scores for all other conditions.

Figure 4.3 Clinical Vignette Scores by Type of Condition

MK Physicians' Scores by Type of Condition
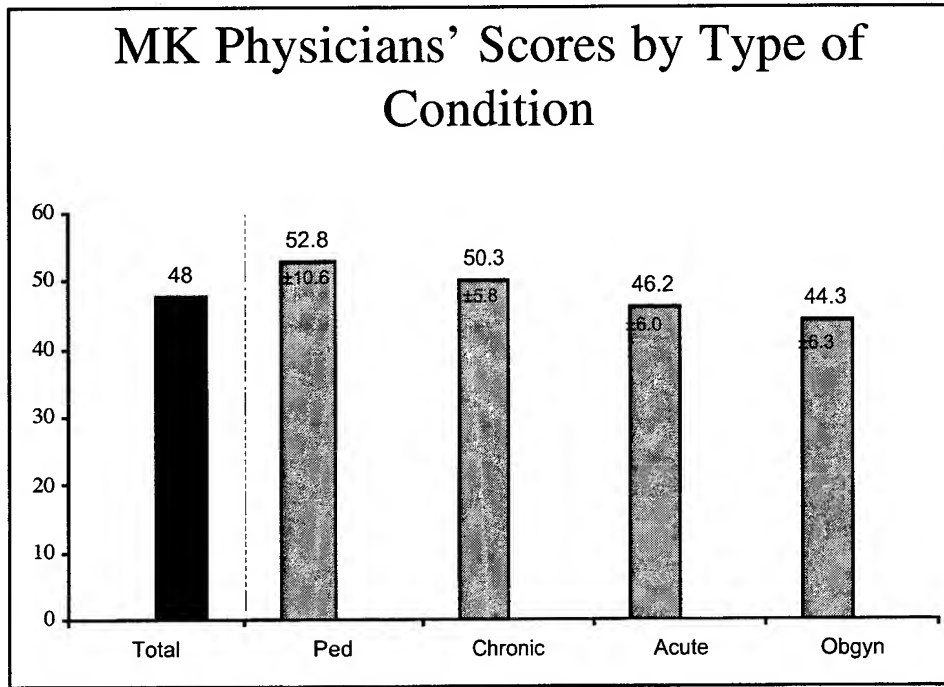
Variation by Group of Conditions. The conditions can be categorized as chronic (BP, CAD, DM), acute (COPD, LB, TB), pediatric (Cough, Diarrhea), and obgyn (Contraception and Prenatal Care). The vignettes scores varied significantly by group of condition (p<0.001).

The highest mean vignette scores was for pediatric conditions (52.8%) and the lowest mean score was for obgyn conditions (44.3%). Two types of conditions (pediatric and chronic) had mean scores above the total mean score of 48.0% and two types (acute and obgyn) had mean scores bellow the overall mean score.

Figure 4.4 Clinical Vignette Scores by Site

Differences by Municipality in Macedonia
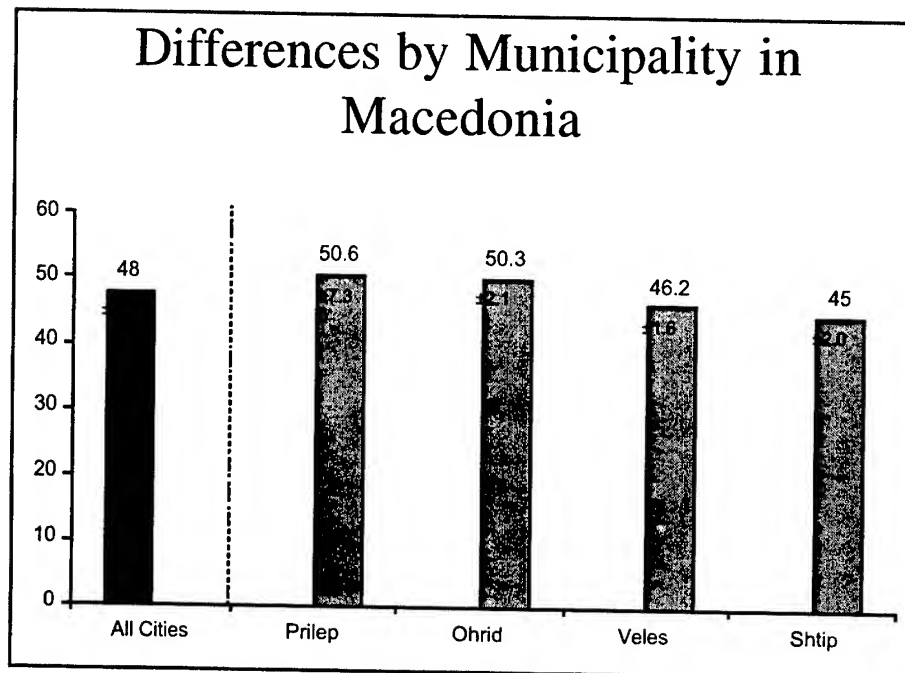
Variation by Municipality. The scores varied significantly (p<0.001) by municipality in Macedonia. Doctors in Prilep received the highest score (50.5%), and the doctors in Shtip scored the lowest average (45.0%).

When compared to the mean of the remaining municipalities, the mean vignette scores for doctors in Ohrid and Prilep are significantly higher, and the mean vignette scores for doctors in Shtip and Veles are significantly lower.

There is however statistical evidence to conclude that there is no difference within two pairs: Ohrid and Prilep, and Shtip and Veles. The mean process score for Ohrid and Prilep (50.4%) is significantly (p < 0.001) higher than the mean process score for Shtip and Veles (45.8%).

Figure 4.5 Clinical Vignette Scores by Age of Physician

MK Physicians' Scores by Age
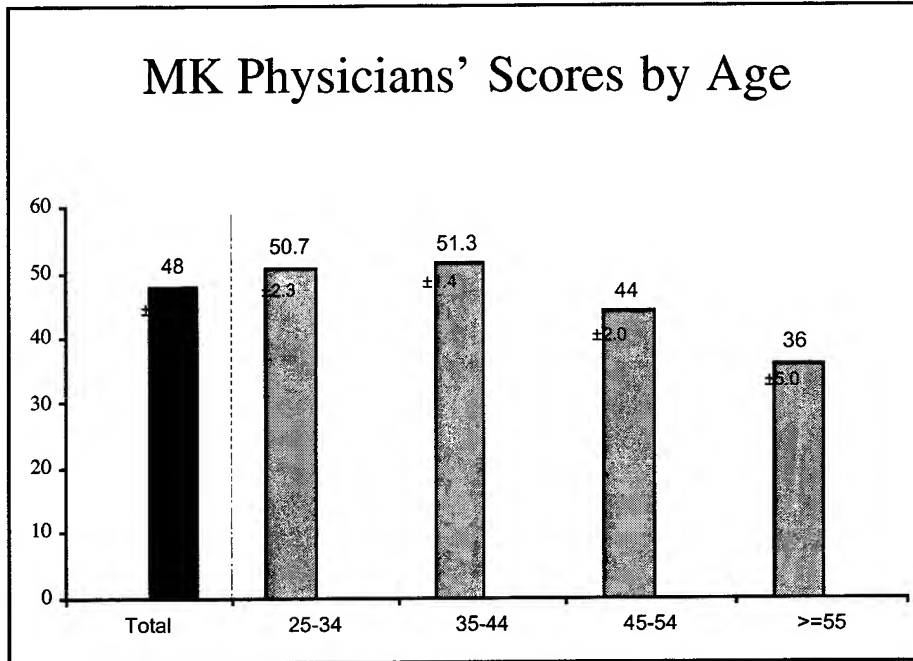
Variation by Physician's Age. Another dimension of variation explored was the age of physicians. The scores varied significantly by (p<0.001) by age, with a decreasing trend in scores as age increases. The highest mean vignette score was achieved by physicians in the age-group 35-44 (51.3%), whereas the physicians with more than 55 years of age had the lowest mean vignettes score as a group (36%).

When compared to the mean of the remaining age categories, the physicians in the 25-34 group and the physician in the 35-44 group scored significantly higher than the mean of the other three respective categories. The physicians in the 45-54 group and the 55+ group scored significantly lower than the mean of the respective three other groups combined.

There is no statistically significant difference between the 25-34 group and the 34-44 group. However, the mean vignette score for doctors in the 45-54 group was statistically higher than the mean vignettes score for the doctors in the 55+ group.

Figure 4.6 Clinical Vignette Scores by Physician Gender
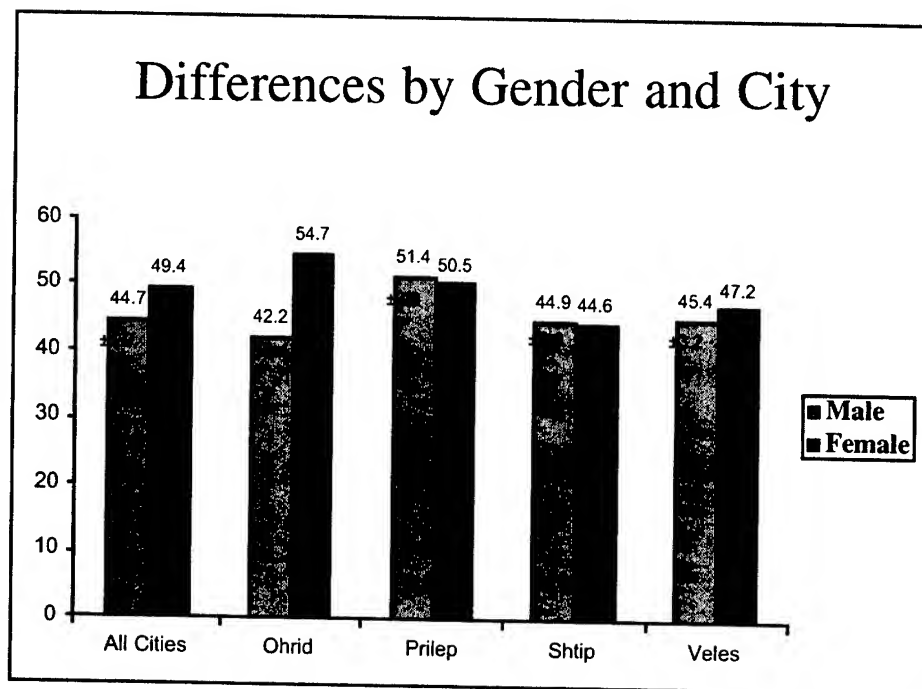
Variation by Physician's Gender. Aggregating the total clinical vignette scores for all cities, female doctors scored higher than their male counterparts. The total score for female doctors is 49.4%, which is statistically significantly higher (p<0.001) than the score for male doctors of 44.7%.

Most of this difference in the descriptive analysis however is explained by the difference in scores in one of the cities: Ohrid. In Ohrid, female doctors scored higher (p<0.001) than male doctors by more than 12 percentage points. In all other cities, the differences between the clinical vignette scores for male doctors and the clinical vignette scores for female doctors were not statistically significant.

Female doctors tend to be younger as well, a fact that may contribute to their higher scores. While the median age category for male doctors is 35-44, the median category for females is 25-34.

Figure 4.7 Clinical Vignette Scores by Physician Specialty

MK Physicians' Scores by Specialty

	School	GP	Labr	Ped	Gyn	ER
Score	50.9	49	48.4	47.2	44.4	43.3
	±2.8	±1.5	±3.5	±2.5	±4.6	±3.9
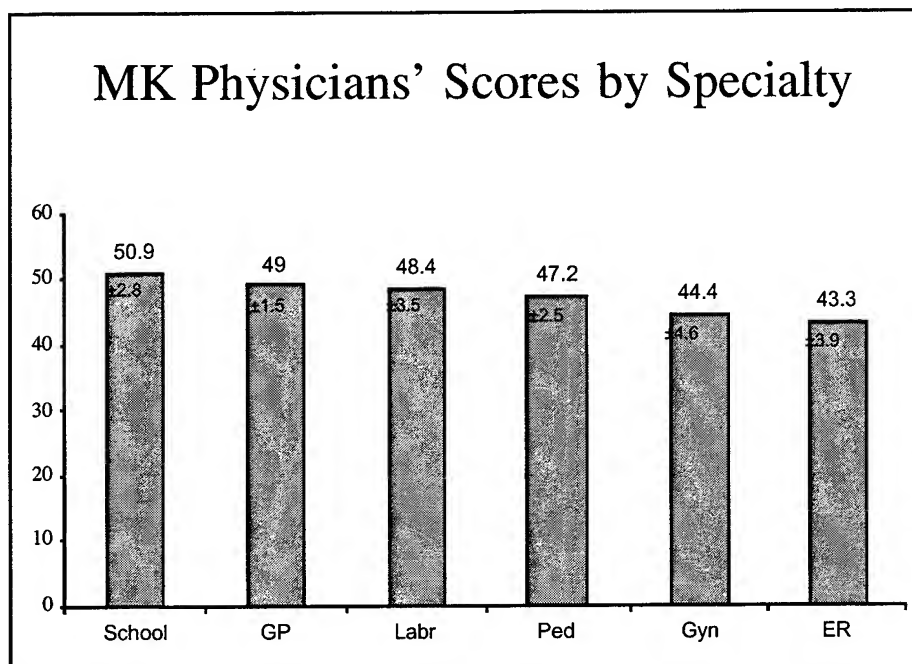
Variation by Type of Practice. The type of practice of physicians in Macedonia also brought significant ($p<0.05$) variation in clinical vignette scores. The different types of practice categorized in the physicians' survey are: general practitioner, pediatrics, obgyn, ER, labor, and school practice.

The highest mean score on the vignettes across all conditions was achieved by physicians practicing in school clinics (50.9%), while the lowest score was that of emergency room doctors (43.3%). Only the mean vignette score for emergency room doctors was significantly ($p<0.01$) different (lower) than the mean of the remaining categories.

Figure 4.8 Clinical Vignette Scores by Domain

MK Differences by Domains

Bar chart showing scores: All Domains = 48, Hx = 56.4, Px = 45.7, Dx = 44.9, Tx = 37.6. Y-axis from 0 to 60.
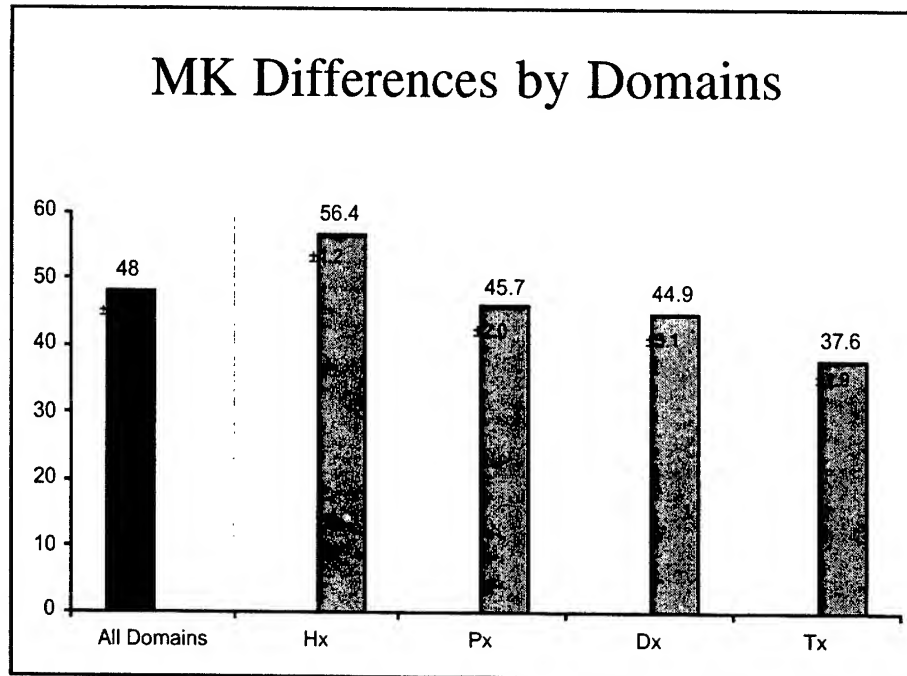
Variation by Domain of Treatment. The domains for which the clinical vignette scores were compared are: History taking (HX), Physical Examination (PX), Diagnostic (DX), and Treatment (TX). This corresponds to the way patients are evaluated in an actual clinical setting in Macedonia and elsewhere and is the reason why vignettes follow this structure.

The scores differ across domains (p < 0.001). The highest mean vignette score is for history taking (57.0%), and the lowest score is for treatment (33.6%). Only the mean score for history taking was above the total mean score.

When compared to the mean vignette scores for the rest of the categories, only diagnostic was not statistically significantly different from the mean of the remaining domains. History taking and physical examination was significantly higher than the average of the

remaining domains, whereas the vignette scores for treatment where significantly lower than the scores for the remaining domains.

Of particular note in the domains is diagnostic, since it carries the weight of correct identification of the condition and serves as a basis for the provision of subsequent care. In 32% of the cases, doctors did not make the correct diagnosis or even a partially correct diagnosis. Only 17% of the doctors received full credit in the diagnostic domain.

A score of zero in diagnostic represents failure to recognize the possibility of the condition (i.e., hypertension for BP, possible TB, etc.). A score of one hundred percent in diagnosis represents the accurate recognition of the condition and its severity. The condition with the highest amount of zero scores on diagnostic was diabetes mellitus, the condition with the lowest was diarrhea. Doctors received a full credit in diagnostic more often in TB and the least times for COPD.

Table 4.2

Comparison of Diagnostic Scores by Condition

Condition[1]	Mean Score for Diagnostic	Percentage of 0% Scores	Percentage of 100% Scores
BP	47%	8%	2%
CAD	63%	14%	39%
Cough	39%	24%	2%
COPD	35%	31%	1%
Diarrhea	61%	4%	20%
DM	17%	76%	9%
LBP	38%	29%	4%
Prenatal	45%	55%	45%
TB	58%	42%	58%

The condition that was most likely to be diagnosed accurately was diarrhea, and the condition that less likely to be correctly diagnosed was diabetes. The condition that was completely diagnosed with 100%

---

[1] There was no diagnostic score for Contraception

accuracy was tuberculosis, and the condition fully diagnosed with the least accuracy was COPD.

A.2. CALIBRATION OF THE VIGNETTES

The vignette scores for Macedonian physicians can be benchmarked with the same instruments used with US doctors. The US vignettes used in this section were validated in clinical practice against actor patients seen in clinics with conditions identical to the clinical presentation described in the vignettes.

The clinical vignette scores for US physicians were close for four of the conditions used in Macedonia: LBP, COPD, DM, and CAD. Those four conditions were identical to the scenarios used for testing Macedonian doctors.[2] The comparison between the scores achieve by US doctors in the four conditions and those achieved by Macedonian doctors in the same scenarios is shown in the figure below.

---

[2] The Total Vignette Score for MK may vary from the aggregate scores observed in the previous section since this comparison uses only four conditions.

Figure 4.9 US and MK Clinical Vignette Scores Across Conditions

Selected MK Conditions
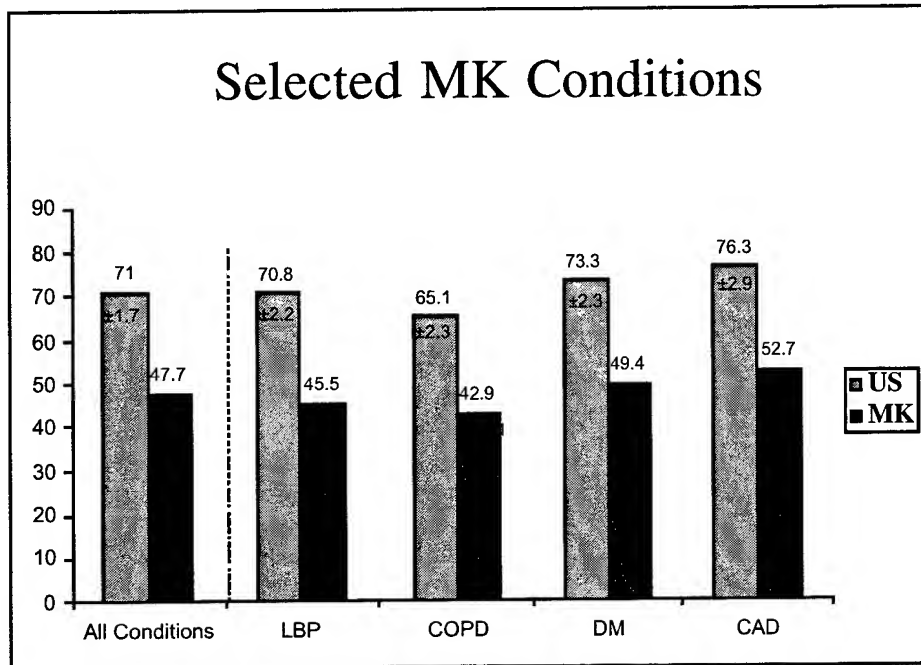
Clinical Vignette Scores Vary Similarly in Macedonia. Figure 4.9 shows the pattern of variation in clinical vignette scores in the US and the corresponding condition score in Macedonia. The overall mean vignette score in Macedonia for the four conditions for which there is a counterpart in the US is 47.7%; this is lower than the mean score for those conditions in the US which is 71%.

The conditions where doctors scored higher and lower are the same in both countries. The condition where the doctors of both countries received the highest score was CAD. The condition where doctors scored lower in both countries was COPD. The total ranking of conditions by score is the same in both countries.

Figure 4.10 US and MK Clinical Vignette Scores Across Domains of Care
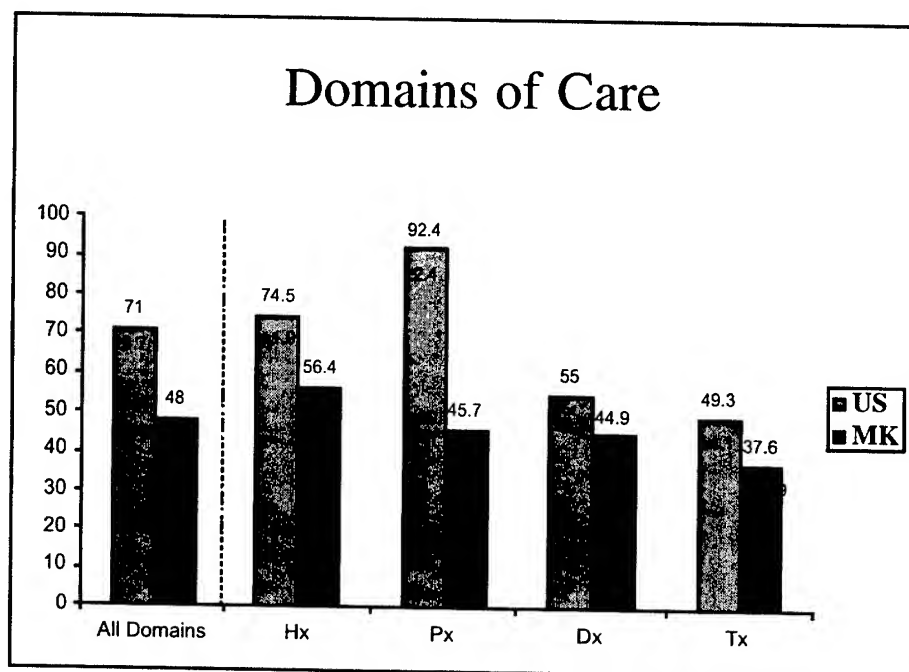
Domains of Care

MK Doctors vary differently with respect to Domain of Care than US Doctors. As illustrated in Figure 4.10, the clinical vignette score extrema for the US and Macedonia don't fully correspond across domains of care.

The domain with the highest score in the US was physical examination, with 92.4%, while the domain with the highest score in Macedonia was history taking, with 56.4%. The lowest scoring domain was the same in both countries: treatment; with scores of 49.3% in the US and 37.6% in Macedonia.

WHAT AFFECTS QUALITY?

So far the analysis has consisted of only univariate analysis of the variability of vignette scores and by extension of the quality of care. Vignettes are a measure of process. Further analysis is required in order to establish relationships between different factors of clinical practice as measured by vignette. Thus, the scores of the clinical vignettes are used as a proxy for quality (process) in the following analyses.

From the results of the analysis of clinical vignette scores against selected descriptive variables, there is strong evidence to suggest that process will vary by site, age of the physician, gender of the physician, and specialty training of the physician. All those variables will be incorporated in the models tested in the following section. The theoretical formulation of the models will also control for other covariates such as workload and type of hospital (public or private) and will follow the theoretical framework described in Chapter 1, where process is a function of structure (i.e. structural inputs).

The first model specification will test the variation in vignette scores using types of condition[3], and controlling for a series of covariates.

Additionally, this research will test the model for each condition individually, and thus not controlling for condition type. This will allow for a better observation of the influences on the process of care score for each specific condition. It is not unreasonable to expect that conditions exhibit different responses to the covariates. A clinic visit for prenatal care may evolve in a different fashion than a visit for diarrhea or tuberculosis.

---

[3] Acute, Chronic, Obgyn, and Pediatric

MULTIVARIATE ANALYSIS OF CLINICAL QUALITY SCORES

B1. IS PROCESS A FUNCTION OF STRUCTURAL CHARACTERISTICS?

In order to explore the different influences that affect the process scores, two categorical measures of structural factors were analyzed: physician-level variables and physical infrastructure of facilities. Following the theoretical framework from Chapter 1 that explains process (the clinical vignette scores) for each condition analyzed as a function of structure (who the physicians are and the conditions under which they practice), we test empirically the nature of the relationship.

Although previous work on health in developing countries assumes that health outcomes variations are a function of structural measures (Rosenzweig 1982, 1983, Rao 1989, Gertler 1990, Evans 1994), the theoretical framework used in this dissertation asserts that the influence of structure on health outcomes is through the workings of process. Thus it is necessary to first how and how much structure affects process before turning to the question of how process affects outcomes.

In addition to the physicians characteristics as explanatory variables, the model also contains measures of physical infrastructure as covariates. Physical infrastructure of facilities (i.e. where process occurs) is considered a measure of quality (Brook 1996, Donabedian 1980, Peabody 1994, 1995, Seidman 1998) so we need to control for different levels of physical infrastructure as well.

The variables available from the facility survey were "number of xx equipment" and "times xx equipment was used," for each possible piece of equipment in the medical facility's inventory. For analysis purposes, however, due to the impossibility of using each and everyone of the physical infrastructure measures separately, the information on

equipment availability and usage was transformed into three different indices of equipment. The three different indices of equipment that were constructed were: sophisticated equipment, basic equipment, and supplies.

Sophisticated equipment comprises that equipment which requires specialized training for its operation and needs skilled technicians for its maintenance: x-ray machines, electrocardiograms, echocardiograms, computers, etc. Basic equipment is that equipment used in primary outpatient visits which can be mostly transported by the physician and is somewhat simple to operate and maintain: stethoscopes, lamps, thermometers, scales, sterilizers, etc. Supplies comprises items that should only be used once and are used commonly in visits to a doctor's clinic, this includes equipment for: suture, intravenous delivery, pregnancy test, masks, syringes, etc. For a rationalization of this specific categorization of medical equipment see Peabody (1994, 1995).

The physical infrastructure measures used for modeling purposes were constructed using data from the facility survey. A correlational scale for both measures was constructed using variables for equipment in each of the categories. Each doctor was thus assigned a score that corresponded to the level of availability and usage of the different equipments comprising a particular category.

The notional model representing the relationship between structural measures and process is the following:

Figure 4.11 Notional Model of Process as a Function of Structure

```
┌─────────────────────┐          ┌──────────────────┐
│      Structure      │          │     Process      │
└─────────────────────┘          └──────────────────┘
┌─────────────────────┐
│ Physician           │
│ Characteristics :   │
│   • Age             │
│   • Experience      │          ┌──────────────────┐
│   • Training        │          │  Events in the   │
│                     │   ───▶   │  Visit to the    │
│ Site:               │          │  Health Care     │
│   • Workload        │          │  Provider        │
│   • Location        │          └──────────────────┘
│                     │
│ Acess to Resources: │
│   • Ownership       │
│   • Equipment       │
│   • Supplies        │
└─────────────────────┘
```
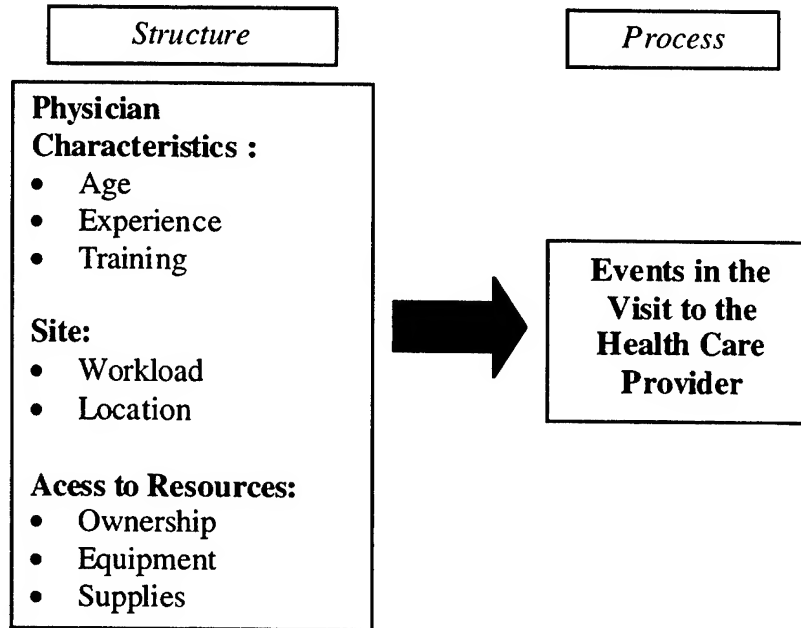
Specifically we tested a model of the form:

Process = f(Clinic Ownership, Physician Characteristics, Physician Workload, Location, Physical Infrastructure of the Facility)

Model 1

In the first multivariate analysis model, the variation of the Process scores (Q) was examined according to 10 explanatory variables:

- Patients Seen Per Week (PP)
- Site (Ohrid & Prilep vs. Shtip & Veles) (S)
- Years of Experience (Y)
- Specialty Training (SP)
- Gender of the Physician (G)
- Hours Worked per Week (H)
- Public or Private Clinic (O)
- Basic Equipment (B)
- Sophisticated Equipment (SO)
- Supplies (SU)

Additionally dummy variables were incorporated into the model to control for the type of condition (Acute, Chronic, Obgyn, Pediatric). Acute was chosen as the default variable. The equation tested was thus:

$$Q = \beta_0 + \beta_1 PP + \beta_2 S + \beta_3 Y + \beta_4 SP + \beta_5 G + \beta_6 H + \beta_7 O + \beta_8 B +$$
$$\beta_9 SO + \beta_{10} SU + \beta_{11} Obgyn + \beta_{12} Ped + \beta_{13} Chronic$$

This equation was tested using simple linear regression and the results are presented in Table 4.3:

Table 4.3

Results of Linear Regression for Model 1

Variable	Coefficient	Robust Std. Err.	t	P>|t|
Public or Private	0.063422	0.482337	0.131	0.895
Patient Panel	1.038742	0.483479	2.148	0.032
Ohrid/Prilep	6.563364	1.144984	5.732	0.000
Years of Experience	-0.653546	0.098785	-6.616	0.000
Specialty Training	4.870036	1.499318	3.248	0.001
Male	-2.460133	1.268906	-1.939	0.053
Hours Worked per Week	-0.194558	0.089843	-2.166	0.031
Public Clinic	0.063422	0.482337	0.131	0.895
Basic Equipment	1.201651	1.099883	1.093	0.275
Sophisticated Equipment	0.512048	1.429773	0.358	0.720
Supplies	2.573551	1.546063	1.665	0.097
Obgyn	-0.598924	1.519966	-0.394	0.694
Pediatric	5.267391	1.920038	2.743	0.006
Chronic	4.233342	1.178601	3.592	0.000

The variables with significant coefficients in this regression model are: Patient Panel, Ohrid/Prilep, Years of Experience, Specialty Training, Hours Worked per Week. The condition type where the vignette scores varied statistically from the scores of acute condition were: pediatric conditions, and chronic conditions.

Patient panel, Ohrid/Prilep, and Specialty training had a positive effect on process. Process of care improved as doctors had more patients, practiced in Ohrid or Prilep, and had received specialty training.

Years of Experience and Hours Worked per Week had a negative effect on process. The longer the doctor had been practicing or how old they were, or the longer workshifts they had, the more likely there were to have lower process scores.

As we say in the univariate analysis, process of care for pediatric and chronic conditions was more likely to be higher than the ones of acute and obgyn conditions.

Model 2

In the second multivariate analysis model, the manner in which the process scores (Q) varied was examined according to the same 10 explanatory variables as in Model 1:

1. Patients Seen Per Week (PP)
2. Site (Ohrid & Prilep vs. Shtip & Veles) (S)
3. Years of Experience (Y)
4. Specialty Training (SP)
5. Gender of the Physician (G)
6. Hours Worked per Week (H)
7. Public or Private Clinic (O)
8. Basic Equipment (B)
9. Sophisticated Equipment (SO)
10. Supplies (SU)

The difference in this case is that instead of using the dummy variables for the type of condition, this analysis ran the same model for each individual disease. Thus, the equation that was tested for each individual condition was:

$$Q = \beta_0 + \beta_1 PP + \beta_2 S + \beta_3 Y + \beta_4 SP + \beta_5 G + \beta_6 H + \beta_7 O + \beta_8 B + \beta_9 SO + \beta_{10} SU$$

Table 4.4 presents the results of the regression of process scores on the covariates outlined in the model.

Table 4.4

Regression Results of Model 2

		Public or Private	Female=0 Male=1	Yrs. Of Experience	Specialist Training	Hrs Wkd Per Week	Ptnts per Week	Ohrid or Prilep	Basic Equipmnt	Sophis Equipment	Supplies	R-sq
BP	Coeff	-1.1850	-0.9016	**-0.7008**	1.6695	-0.5878	1.3220	5.4909	-1.7170	2.0839	5.1358	0.1826
	se	1.9144	6.7284	0.3872	7.3546	0.4592	1.8184	3.8211	2.7605	5.0458	3.7460	
CAD	Coeff	-0.7150	-1.9660	**-0.4414**	4.1086	-0.3743	-0.1770	4.0608	-1.3980	6.5291	-4.4640	0.1658
	se	1.2478	3.2931	0.1962	2.7012	0.3199	1.4874	3.1700	2.9370	4.0353	5.4439	
Cough	Coeff	1.6195	-2.8870	**-1.1370**	6.7207	-0.2235	0.1829	**9.8777**	4.3109	-6.2770	6.0717	0.3489
	se	2.3471	3.8340	0.4323	6.3764	0.3292	1.7931	4.6374	4.0823	5.3195	6.2984	
COPD	Coeff	1.2262	2.0201	-0.2034	-2.1260	-0.1147	2.2396	_9.7826_	3.5771	**-7.9090**	4.5625	0.3453
	se	1.1802	4.2262	0.2150	3.0071	0.1821	1.1695	2.9272	2.0382	3.5466	3.5015	
Contr	Coeff	1.3404	-3.8930	**-0.6512**	0.3104	-0.1244	1.5630	5.8960	2.3381	0.5528	3.3056	0.1676
	se	1.5044	3.5072	0.3101	5.1733	0.3248	1.9178	3.9280	4.3867	5.3670	8.9341	
Diar	Coeff	1.6072	**-14.3200**	-0.9888	12.3170	0.0065	2.4677	**16.0940**	9.0561	-_14.6500_	3.7491	0.6527
	se	2.1147	-4.2367	0.5851	9.8932	0.6500	1.3574	4.8903	4.8068	5.1224	2.9850	
DM	Coeff	-1.0630	-0.7282	**-0.6667**	5.6262	0.0287	1.1809	_10.0230_	0.6636	1.8545	0.8024	0.2245
	se	1.3490	3.8529	0.2058	3.6158	0.2158	1.4401	3.5656	2.9105	4.6831	4.5079	
LBP	Coeff	0.1106	0.8370	**-0.7701**	8.1890	-0.3267	-0.5315	4.7748	0.4477	5.5379	-3.9400	0.2152
	se	1.2860	3.0886	-0.2567	4.6528	-0.1785	-1.1605	2.5398	4.2638	4.0130	4.1693	
Prena	Coeff	-1.8260	-7.3880	**-0.7073**	8.8123	-0.0662	1.4073	-1.2440	0.6632	9.8132	7.4228	0.2161
	se	1.6752	4.2953	0.3062	4.7074	0.3719	2.2197	4.6245	-3.7258	5.3919	7.0425	
TB	Coeff	-0.3583	-0.9085	**-0.6673**	4.1536	-0.0016	0.0388	5.3204	3.3862	-8.8260	**10.0250**	0.2881
	se	1.3077	3.7234	0.2254	3.8459	0.1778	1.1086	3.6667	4.2065	4.6947	3.0657	

Highlighted: p ≤ 0.10   Bold: p ≤ 0.05   Underlined: p ≤ 0.01

Public or Private Clinic. There were no statistically significant differences between the different types of provider for any of the conditions.

Clinical Quality for physicians working in public or private clinics was equally distributed and there is no statistically significant evidence to support any differences between them.

Physicians. There are statistically significant differences among physicians relating to their years of experience. Across all conditions with the exception of COPD, physicians who were more recent graduates score higher than physicians with more years of experience[4].

Specialist training contributed to higher clinical quality for LBP and Prenatal Care. Male physicians tended to provide lower clinical quality than their female counterparts, but that relationship was statistically significant only for Diarrhea and Prenatal Care.

Physician Workload. The number of patients seen in a week contributed to higher clinical quality for COPD and Diarrhea, the more patients a physician examined, the higher the clinical quality score.

Hours worked per week was statistically significant for LBP only, it contributed negatively. More hours worked per week were translated to lower clinical quality but only for low back pain.

Location. The location (municipality) of the physician had a statistically significant impact on clinical quality. Based on the results of the descriptive statistics, two municipality "groups" were constructed: Ohrid and Prilep, and Shtip and Veles.

---

[4] Different model specifications were tested in order to discern whether the relationship of years of practice was nonlinear, but this testing didn't yield any significantly different results.

Polynomial formulations on years of experience were tested with the same results.

Being located in Ohrid or Prilep contributed significatively and positively to the physician's vignette scores when compared to the physicians in Shtip or Veles for Cough, COPD, Diarrhea, DM, and LBP. Only the coefficient in the Prenatal Care regression had a negative sign, but that relationship was not statistically significant at the 0.10 level.

Infrastructure Measures. Basic Equipment had a positive influence and was statistically significant for only COPD and Diarrhea. Sophisticated equipment was statistically significant but with a *negative* coefficient for COPD, Diarrhea, and TB, but a positive coefficient for Prenatal Care. Supplies had a statistically significant contribution for TB only, with a positive sign.

Years of experience and site were the variables that more consistently explained variations in clinical quality. Model 2 provided a better explanation of variations in clinical quality of diarrhea.

B2. USING CONDITION SPECIFIC EQUIPMENT

To better elucidate the conflicting effects of equipment and supplies on process, the same equation from Model 2 was tested for each individual condition using condition specific equipment scales. These scales were constructed by looking at the equipment (basic and sophisticated) and supplies that are specifically needed in the clinical care of an individual condition. After consulting with primary care physicians, the equipment contained in the facility survey, I identified the different equipment and conditions. A correlational scale was constructed for each condition using the availability of the equipment and supplies identified as relevant to the treatment of the condition. Each doctor was thus assigned an individual score per condition, conditional on the availability of the condition-specific items in the doctor's facility.

A likelihood ratio test was constructed using the condition-specific regression model with the equipment index and without the equipment index. For the majority of the conditions, the addition of a condition-specific equipment index did not make a difference, i.e. there were no improvements in model's level of determination (R^2) or significance (F).

Table 4.5

Results of Tests of Model 2 with Condition Specific Equipment

Condition	χ^2	p - value	Does Equipment Improve the Model?
Blood Pressure	0.13	0.72	No
CAD	3.64	0.06	No
Cough	0.02	0.90	No
COPD	0.68	0.41	No
Contraception	0.40	0.53	No
Diarrhea	19.0	p<0.001	Yes
Diabetes Mellitus	1.55	0.21	No
Low Back Pain	0.85	0.35	No
Prenatal Care	7.25	0.01	Yes
Tuberculosis	4.10	0.04	Yes

The addition of condition-specific equipment index, however, improved the model for three conditions: Diarrhea, Prenatal Care, and Tuberculosis. In all cases the coefficient for equipment was positive.

The multiple regression models with the condition-specific equipment index followed the same pattern as the regression models with different equipment levels tested in B1. The variables that were more consistent in explaining the behavior of clinical quality were the site and years since graduation. They were not only consistent in their statistical significance, but also in the sign observed in the previous tests.

Table 4.6

Significance and Sign of Parameters for Years of Experience and Site in Model 2 with Condition Specific Equipment

| Condition | With Equipment Scale | | | | Without Equipment Scale | | | |
| | Years of Experience | | Ohrid & Prilep | | Years of Experience | | Ohrid & Prilep | |
	p-value	Sign	p-value	Sign	p-value	Sign	p-value	Sign
Blood Pressure	0.07	–	0.17	+	0.06	–	0.10	+
CAD	0.01	–	0.27	+	0.03	–	0.05	+
Cough	0.01	–	0.04	+	0.01	–	0.02	+
COPD	0.49	–	0.01	+	0.44	–	0.00	+
Contraception	0.04	–	0.10	+	0.03	–	0.10	+
Diarrhea	0.03	–	0.00	+	0.06	–	0.01	+
Diabetes Mellitus	0.00	–	0.00	+	0.00	–	0.00	+
Low Back Pain	0.01	–	0.01	+	0.00	–	0.01	+
Prenatal Care	0.04	–	0.27	+	0.02	–	0.21	+
Tuberculosis	0.01	–	0.61	+	0.00	–	0.13	+

The figures in Table 4.6 suggest that years since graduation and site are consistently strong predictors of clinical quality. Years since graduation, as was observed in Model 1, has a negative coefficient. This means that as doctors are further distanced from their medical qualification training, their knowledge of updated clinically accepted practices suffers.

Site is also a strong predictor. A careful analysis of the differences and similarities of the different municipalities may allow

elucidation of the factors that are encompassed in the site variable. Just as it was observed in Model 1, the clinical quality scores in Ohrid and Prilep are consistently higher than those in Shtip and Veles.

IS PROCESS RELATED TO OUTCOMES?

The importance of the analysis of quality has to be in the impacts that we can expect on actual patients well being. Although conceptually the idea that improvements in different areas of quality of care will improve health status in a population is often assumed, the contribution of this dissertation is to provide some of the first empirical evidence for this relationship. Even though the linkage between process and health outcomes is often assumed without question and is the basis of national health policy and budget allocations around the world, it is an area that is lacking in empirical studies that test this relationship. To this end, the clinical quality scores of the primary health providers will be tested against the health outcomes of the population being served by them.

The first of the models developed in this section will test the effect of the process of care scores of physicians for a particular health facility against the self-assessed health status of the patients living in that area that use or at least can potentially use the facility. The linking of the scores was conducted through the codes in the exit and household health surveys. One of the strengths of the study design, is that it uses a census of health facilities and health-care providers, therefore the analysis can operate under the premise that we are evaluating the universe of health care providers in a community and thus any person within that community will receive her care in the respective facility.

Based on a census of health-care providers in a community, the health outcomes of patients in the vicinity of a health facility are compared with the process of care scores for the doctors in a particular facility within the municipalities.

C.1. DOES PROCESS AFFECT OUTCOME?

The fundamental question that has to be examined after the results of the analysis of process of care scores is whether the process scores are related to measures of health. Based on the 1800 household interviews of "exited" patients, and controlling for the patients' individual characteristics, we looked at the impact of clinical quality on self reported (subjective) health scores:

Figure 4.12 Notional Model of Health Outcomes as a Function of Process

Process		*Outcome*

Events in the Visit to the Health Care Provider	➡	**- Self-Assessed Health - Clinical Indicators**
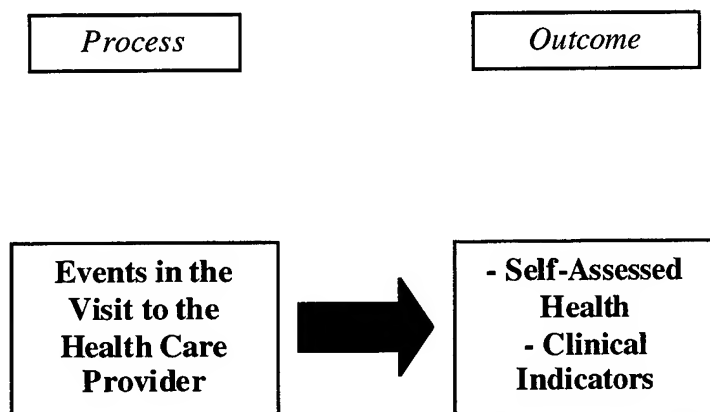
For the purposes of testing the incidence of process of care on health outcomes, the model looks as follows:

$$\text{Self Assessed Health} = f(\text{Process Scores, Age, Gender, Education, Income, Access to Information})$$

The measure of health used in this model is subjective. The self reported health assessment from the exit survey was used as a proxy for health. The question read "How do you rate your health?" and the responses varied on a 5 level scale from excellent to very poor. The possible answers are: Excellent, Very Good, Good, Fair, Poor. For the purposes of this analysis, "Poor" was given a score of 1 and Excellent a score of 5, in order to associate better health with a higher number in the dependent variable.

Model 3

Because of the nature of the responses, an ordered logit model was used to test the relationship. Formally, what we are assuming is that the probability that a person will rate her health more favorable can be approximated by:

$$\text{Prob}(\Delta H > 0) = \Phi(Q, A, G, E, Y, N, T, O)$$

Where:

H: Health

Q: Clinical Quality of Care

A: Age

G: Gender

E: Education

Y: Income

N: Access to Information

T: Ethnicity

O: Clinic Ownership[5]

The variables used to test Model 3 were:

- Self Assessed Health (1 -5 Scale)
- Average Clinical Quality Score for doctors in the Patient's Community
- Age of Patient
- Gender (Female=3, Male=1)
- Number of School years Completed
- Income
- How often the patient reads the Newspaper
- The language spoken at home (1=Macedonian, 0=All Others)
- Type of Clinic Public vs. Private

---

[5] Although not a demographic variable, this is used to control for selection of patients that visit different types (public vs. private) of clinics.

This model was tested using maximum likelihood estimators. Table 4.7 presents the results of the regression.

Table 4.7

Results of Ordered Logit Regression of Model 3

Variable	Coefficient	Robust Std. Err.	z	P>|z|
Process Score	0.2406282	0.0408099	5.90	0.000
Age	-0.0640275	0.0073276	-8.74	0.000
Female	-0.5024785	0.1088140	-4.62	0.000
Income	6.61e-07	3.12e-07	2.12	0.034
Education	0.0310616	0.0713726	0.44	0.663
Reading News	-0.0507222	0.1041800	-0.49	0.626
Macedonian	-0.2774206	0.3609179	-0.77	0.442
Public or Private	0.2773127	0.2032355	1.36	0.172

Process Score. The clinical vignette scores were aggregated by facilities within cities. The number of facilities varied by municipality as presented in Chapter 3. We are assuming a group of patients that is examined by a group of doctors. We are also assuming that patients see doctors in a particular facility within their municipality.

The process scores had a positive effect on health status. The higher the average process of care scores of the group of physicians in the community, the higher the health score of the patients ($p < 0.001$). This suggests that the knowledge of physicians in a given community has an impact on the health of the population.

Age. Age is negatively correlated with subjective report of health. The older a person is the less likely that person is to report that she feels healthy ($p < 0.001$).

Gender. Gender has a significant effect on health. Women rated their health lower than men ($p < 0.001$).

Income.   Income presented a statistically significant positive effect on self assessed health (p>0.05) when controlling for other covariates.   The higher the income of a person, the more likely the person is to report feeling well.

Education.   Education, as measured by highest school year completed, had no statistically significant impact on health.   Although the coefficient was positive, it had no statistically significant effect.

Reading News.   People who regularly read the news were no more likely to report a higher score in their health assessment than people who did not read the news regularly.

Ethnicity.   Using the language spoken at home as a proxy for ethnicity, we used a binary variable to identify those who speak Macedonian at home versus other languages.   Ethnicity did not provide evidence of affecting self reported health status.

Clinic Ownership. The type of clinic visited by the patient (public or private) had no statistically significant impact on self reported health status.

The results of the model strongly associate process quality and health.   The control variables strongly strengthen this conclusion that coincides with the expected results based on the theoretical framework and the existing literature.   Overall the model presents statistically significance (p<0.001) and its coefficient of determination is adequate for aggregate level data (0.18).

C.2. DOES PROCESS RELATE TO OBJECTIVE MEASURES?

Although the subjective measure of health used emanates from widely validated instruments (Hays 1990, Keller 1998, Litwin 1998), subjective reports are well known to be biased (Ford 2001, Robertson 2001, Miller 2001). Thus, in order to strengthen the findings of the previous analysis, we tested the relationship of process scores with objective health measures.

For the purposes of testing the relationship between process scores and objective health measures, three measures were selected: blood pressure, blood glucose content, and lung capacity. These objective measures where then compared with the scores on clinical vignettes for corresponding scenarios in a medical-facility based linkage. Blood pressure for a patient was compared with the average process of care score for the blood pressure vignette for the doctors in the health facility in the geographical area where the patient was seen. Glucose levels and lung capacity were compared with the average process of care scores for DM and COPD respectively.

Table 4.8 presents the variation of the objective health measures across cities, there are statistically significant differences between the cities, thus there is a need to control for municipality.

Table 4.8

Mean Objective Measures of Health for Healthcare Facility Users

City	Blood Pressure [1]	Blood-Glucose [2]	Lung Capacity [3]
Ohrid	82.6	110.3	228.0
Prilep	85.8	166.8	288.7
Shtip	84.9	103.1	304.9
Veles	81.4	162.2	320.0

1. No statistically significant differences in means or variance between cities
2. Statistically significant differences in means and variance between cities ($p < 0.01$)
3. Statistically significant difference in means only between cities ($p < 0.01$)

As a preliminary comparison between the subjective-based vs. the objective based model, we compared subjective to objective measures of health status. A Spearman-Rank correlation test rejected the null hypothesis that objective and subjective measures were independent at the $p<0.01$ level. Our self-reported health status has an actual relationship to the way our bodies are functioning. This comparison further validates the use of subjective measures in our study.

The comparison of process scores with the objective health measures enriches the analysis, however, and allows us to explore the workings of process in a more direct way. Since the objective measures do not conform to a single scale, the relationship between process scores and objective measures was explored with a logit model for the binary choice good health vs. bad health[6]. The logit model tested the variation in good health/bad health against the variation in the mean condition specific vignette scores for physicians by health facility, controlling for age and gender.

_____

[6] good health: systolic pressure < 140 and diastolic pressure <90, blood glucose content < 127, lung capacity >250

Model 4

In order to test the relationship of process of care scores on objective health outcomes, the model we used tested the probability that the clinical indicators of health status (good vs. bad) was dependent on the process of care scores of the physicians serving in the patient's community.

In this case, we tested the relationship of process scores and the probability of good objective health, controlling for covariates that affect the patient's health such as age, gender, income, education, access to information, and type of clinic visited. Given the descriptive statistics of objective outcomes presented in Table 4.6, the model also controls for the Municipality where the patient lives.

The probability that a patient health will be good or bad depending on the process of care scores of the doctors in the patient's community, can be approximated by:

$$\text{Prob}(H) = \Phi(Q, A, G, E, Y, N, O, M)$$

Where

H: Health Status

Q: Clinical Quality of Care

A: Age

G: Gender

E: Education

Y: Income

N: Access to Information

O: Clinic Ownership

M: Municipality

The variables used to test Model 4 were:

- Good or Bad Health based on three outcomes: blood pressure, blood glucose content, and lung capacity

- Condition Specific Average Clinical Quality Score for doctors the Patient's Community

- Age of Patient

- Gender (Female=3, Male=1)

- Number of School Years Completed

- Income

- How often the patient reads the newspaper

- Type of Clinic (Public vrs. Private)

- Three dummy variables to identify the cities (Prilep is the default)

Table 4.9

Results of Logit Regression of Model 4

Variable	Coefficient	Robust Std. Err.	z	P>|z|
Process Score	9.2976500	1.9198680	4.84	0.000
Age	-0.0534605	0.0083876	-6.37	0.000
Female	-0.2741704	0.1181236	-2.32	0.020
Income	-1.64e-07	2.23e-07	-0.73	0.463
Education	0.0431676	0.0774627	0.56	0.577
Reading News	0.2347314	0.1231245	1.91	0.057
Public or Private	0.0540375	0.0850833	0.64	0.525
Veles	1.7007920	0.4531899	3.75	0.000
Ohrid	0.5803789	0.2710769	2.14	0.032
Shtip	1.2068160	0.3335050	3.62	0.000

Process Score. The results of logit model suggests that the condition specific process scores contribute significantly to better health outcomes (p<0.001) when controlling for covariates of the patient. The confidence interval of the odds ratio for the process scores does not contain 1, indicating a statistically significant and a positive change on the probability of better health. The observed coefficient for process score would indicate that a unit increase in process score will beget a 2% probability increase that the patients will have good health.[7]

---

[7] In a logit model, the rate of change in the probability is given by $\beta_j P_i (1-P_i) \cong 0.25 \ \beta_j$ (Amemiya 1981)

Age. Age is negatively correlated with subjective report of health. The older a person is the less likely it is that the person will enjoy good health (p < 0.001).

Gender. Gender has a significant effect on health. Being a woman decreases the probability of enjoying good health (p < 0.05). Bearing and raising children, together with other disadvantages that women have in many societies, may take its toll on the health status of women.

Income. Income presented no statistically significant effect on health status (p>0.05) when controlling for other covariates. The income distribution observed in Macedonia does not vary much, unlike that observed in other countries with a longer tradition of market economies. This, coupled with the public provision of health care, may be the cause for the lack of effect of income on health status.

Education. Education, as measured by highest school year completed, had no statistically significant impact on health. Although the coefficient was positive, it had no statistically significant effect.

Reading News. People who regularly read the news were no more likely to report a higher score in their health assessment than people who did not read the news regularly.

Clinic Ownership. The type of clinic visited by the patient (public or private) had no statistically significant impact on the probability of enjoying good health.

Municipality. The Municipality where the patient lives has a statistically significant impact on the patient's objective health. All Municipalities differed from each other (p<0.05), providing evidence of a strong location effect.

The regression results suggest that in 74% of the observations it was possible to predict the health status correctly based on the values of the process of care scores and controlling for other covariates in this model. Overall the model presents statistical significance (p<0.001) and an adequate coefficient of determination for aggregate level data (0.15). As expected from our theoretical framework, the process score had a positive contribution on objective health outcomes.

TABLE OF CONTENTS FOR CHAPTER 5

DISCUSSION

Developing countries are believed to have lower quality of care compared to developed nations (Gani 1996, Omaswa 1997, Pécoul 1999, Seidman 1998, Shawyer 1996, Thaver 1998), yet there are few quantifiable measures of quality of clinical care in developing countries (Durán-Arenas 1998, Ziv 1998). One of the reasons may be that direct measurement of clinical process and outcomes is complicated by factors including case mix variation, different clinical facilities, and perceived standards of care. This confounds the process of formulating adequate policies that address the public health needs of developing countries since it is difficult to have an adequate diagnosis of the health care system (Bhat 1996, 1999).

Sources of variation in medical practice are attributable to the availability of resources-structure-and the inadequate application of medical knowledge and science-process-(Kalf 1996); both sources have been quantified in this study. The present study explored variations in the area of medical knowledge first, and then examined the relationship of the knowledge levels with resource availability, and most significantly, to health outcomes. This was done approaching the variations from a quality of care perspective.

The different quality of care measures as suggested by Donabedian (1980) are:
- Structural Measures—physician characteristics and physical means (i.e. hospital, equipment)
- Process Measures—the factors of the encounter between the health care provider and the patient
- Outcome Measures—the subsequent health status of the patient.

Process measures are preferred as tools for assessing quality since they allow for the identification of acceptable and appropriate care of a patient. Outcome measures can depend on spurious relationships that

don't reflect the level of care received by the patient (Brook et al. 1996). Health outcome measures also present evidence of confounding factors such as comorbidities and socioeconomic influences on health—elements that are regularly outside of the sphere of dominion of the health provider's daily practice.

However adequate process-based quality of care measures might be, for them to be valid there must be evidence that changes in the process result in changes in patients' outcomes (Brook et al. 1996). In the case of the clinical vignettes used in this study, there is ample empirical support for the effects of the measured process on patient's health since the scoring of clinical vignettes is grounded on evidence-based medicine and the vignettes have been thoroughly validated.

Although the extrinsic sources of variation in medical practice can be availability of resources or inadequate application of medical or scientific knowledge (Kalf 1996), the debate on improving quality of care and eliminating differentials between regions draws heavily on the resource and financing aspects (Seidman 1998, Peabody 1994, Pécoul 1999) or on the regulatory conditions (Bhat 1996, 1999). This study directly examined the dimension of medical and scientific knowledge and practice, and how the variations manifest themselves with respect to inputs and outcomes.

The research framework of this study, as illustrated in Figure 5.1, considered the effects of structural measures on health outcomes as the workings of process, and thus the theoretical relationship that links the three dimensions of quality assumes a critical role for process: it is the catalyzer through which the inputs into the production of health by individuals and communities are transformed to generate health outcomes and indicators.

Figure 5.1 Theoretical Framework of this Study

Structure	Process	Outcome
Physician Characteristics : • A g e • E xperience • T raining **Site:** • W orkload • L ocation **Acess to Resources:** • O w n ership • E quipment • S upplies	**Events in the Visit to the Health Care Provider**	**- Self-Assessed Health** **- Clinical Indicators**
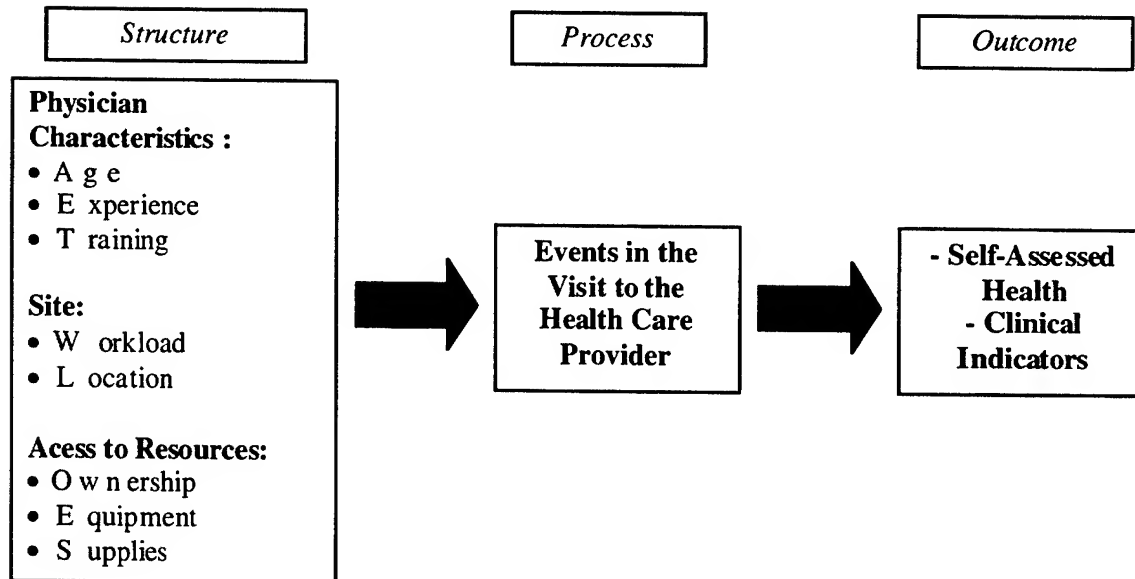
This study relied on primary data collected in Macedonia in 1997 to measure the quality of health care in a less developed country. We explored variations in the area of medical knowledge first (a process measure) for nine common health conditions that are best treated in a primary care setting: coronary artery disease (CAD), hypertension (BP), chronic obstructive pulmonary disease (COPD), diabetes (DM), prenatal care (Prena), tuberculosis (TB), contraception (Contr), and low back pain (LBP) for adult patients, and diarrhea (Diarr) and cough with fever (Cough) for children. Following the descriptive analysis of the process measure, we examined the relationship of process with structural aspects of health care, including resource availability. A final analysis examined the relationship of process with health outcomes at the population level. This was done approaching the analysis from a quality of care perspective.

RESULTS

The health outcomes were analyzed against the vignette scores using the following notional model:

$$\text{Health} = f(\text{Process Scores, Sociodemographic Covariates})$$

This notional model was analyzed with two different sets of dependent data:

- using subjective health measures—a five point scale answer to the question: how do you rate your health?
- objective health measures—blood sugar content, blood pressure, and pulmonary capacity.

As the analysis of health outcomes presented in Chapter 4 showed, the process measures used in this study are strongly associated with subjective and objective health outcomes, this supports the idea that the differential in medical and scientific knowledge matters as an indication of the source of variations in health status. Controlling for sociodemographic covariates that affect health, the higher the average scores on clinical vignettes for a group of doctors in a geographically classified area, the more likely it was that the population in that geographic area enjoyed good health—both self assessed and measured through objective health indicators. These findings open up the possibility of discerning a policy lever in the quest of improving the health status in developing countries: medical knowledge.

The descriptive analysis of the clinical vignettes scores showed that there was significant variation in the process of care across different dimensions of interest. The statistically significant variation in clinical vignette scores by condition and by type of condition suggest the possibility of knowledge variations across diseases. Given that the clinical vignettes represent conditions

commonly found in outpatient settings, there is no *a priori* information to suggest that the variation should exist.

The variation by municipality is consistent with previous studies on the diffusion of medical knowledge (Phelps 1992, Dosi 1988) and in general with geographic variation in medical practice (Brook 1984, Chassin 1987, Berwick 1991, Eddy 1984, Weiner 1995). This is one reason to control for municipality in the multivariate analysis. The statistically significant differences by gender and age were taken into account when performing the multivariate analysis.

The multivariate analysis of the influences on process was performed using the clinical vignette scores as the dependent variable in the following model:

$$\text{Process} = f(\text{Clinic Ownership, Physician Characteristics, Physician Workload,}$$
$$\text{Location, Physical Infrastructure of the Facility)}$$

This relationship was tested using two models:

Model 1: Using a dummy variable for the type of condition[1]
Model 2: Running the model for each condition individually

In the multivariate analysis of the influences on process, after controlling for other structural covariates affecting medical knowledge, the years elapsed since the graduation from medical training is a main influence across conditions. A higher number of years since finalization of formal medical training consistently contributed to lower scores on the clinical vignettes. The interpretation of this finding is that the more recently graduated physicians are more current on medical knowledge because they have been exposed to the latest developments in their training. This highlights the fact that practice in itself is not the contributing factor to better medical knowledge, but practice that is accompanied with an actualization of knowledge. In interpreting this finding, it is assumed that there are no deterioration

---

[1] Conditions were classified in the following types: Acute, Chronic, Obgyn, Pediatric

effects in doctor's knowledge associated with age since 92% of the doctors in the study were younger than 55. Thus the effects observed are more likely to be associated with the time that has elapsed since the doctor left formal medical training. In trying to formulate policy influenced by this finding, interventions geared towards actualizing clinical knowledge thus become strong candidates for the recommendations to be offered in order to improve a health system.

The results of Model 1 and Model 2 do not provide clear evidence regarding the relationship between structure and process. These results have to be approached with caution though, since it is reasonable to believe that the three measures of infrastructure are related, thus creating a problem of multicollinearity in the model. Multicollinearity will not allow for the proper estimation of the standard errors and therefore lead to biased estimation of the coefficients and inaccurate conclusions regarding the level of statistical significance (or lack thereof) of the coefficients. The coefficients however remain Best Linear Unbiased Estimates (BLUE), and therefore their signs can be considered reliable.

We measured p to determine if the infrastructure measures were correlated. This analysis revealed a correlation of 0.53 between the supply index and the sophisticated equipment index. Eliminating one of the indices from the model can lead to misspecification bias. Since the significance levels and coefficients of the other covariates do not change from the previous model on structural measures, this model specification remains the preferred one.

Using condition-specific equipment indices, three out of the ten conditions exhibited a benefit in a model formulation that included equipment. These conditions were diarrhea, prenatal care, and tuberculosis. This finding further supports the concept that the contribution of infrastructural measures is condition specific and is more suitable for addressing a specific need.

Through the analysis of the clinical vignette scores, this study provides evidence of the differential in quality of care in the Former Yugoslav Republic of Macedonia and the United States. It also confirmed the influence of practice conditions and medical education on the ability for applying medical knowledge to the interaction with patients, which translates into better health outcomes.

The costs of most infrastructure measures—introducing a national drug supply, constructing a hospital, equipping a facility—seem prohibitive compared to the costs of continuing medical education. In addition to the higher costs, this paper suggests that an investment in medical education can yield better overall health outcomes than improvements in infrastructure.

LIMITATIONS OF THE STUDY

BEST WAY TO MEASURE PROCESS?

This study presented evidence of the contribution of different factors to the health status of the population. However, the tools used in measuring the factors (especially process) are not the only existing ones. The literature on the options to measure the process of care indicates that standardized patients (SPs) are the realistic gold standard to evaluate the events occurring in a clinic visit (Rethans 1987, Colliver 1993, Pieters 1994, Badger 1995, De Champlain 1997, Colliver 1995, Glassman 2000).

Three ways could clearly be considered effective in creating a "Gold Standard" of process of care:

1. Double Masked Evaluations of Provider and Patients
2. Recorded Sessions
3. Standardized "Actor" Patients

Option 1. is not feasible to conduct due to human subjects' considerations and ethical reasons, as well as logistical impossibilities of coordinating and monitoring the encounters. A recorded session has the drawback of a high probability of a Hawthorne effect on the part of the physician. The SP is the most feasible approach in the US. It also has the advantage of controlling for case mix variation.

The disadvantage of using SPs is the high cost and complicated logistical procedures that are required for a true one-sided blinded study[2]. The logistical procedures also make each SP visit a rather

---

[2] In order for the clinic visit to be a realistic event, it is necessary to have access to hospital records and create a patient's "history" with varying degrees of sophistication depending on the

expensive event, limiting thus the number of such episodes that any particular study is capable of affording.

Clinical Vignettes have been validated against the gold standard (Peabody 2000, Dresselhaus 2000), they have proven to be reliable measures of process of care and allow for case mix control. Vignettes have only been validated in the US however, and not in other settings.

This is a limitation of the reach of the study. This study represents the first time that these vignettes are used in a setting outside the United States. The uniqueness of this study (an innovative use of clinical vignettes) is also one of its weaknesses. This opens up the possibility for further studies though, since the use of vignettes provides with an economically feasible way of gauging process.

HOW GOOD IS THE SCORING?

The Clinical Vignettes used in this study are paper case scenarios, which are susceptible to problems of readability and interpretation. The proper way to address these shortcomings of the clinical vignettes is through the development of computer-based scenarios that allow for a more faithful read of the physician's responses. A computer-based scenario in turn might contain some bias against older generations of physicians who are not used to the intensive use of data processors that younger generations might have fully incorporated into their training and practice.

ARE THE US-MK RESULTS COMPARABLE?

The clinical vignettes used in this study were developed and validated in US hospitals. As mentioned before the vignettes had never been validated with actor patients outside of the US prior to this

hospital's records system and the case. The creation of prior laboratory tests, x-rays, and other medical history might be necessary.

study, so this is a factor that needs to be taken into consideration in interpreting the results.

Further exploration is needed on how clinical vignette scores in different settings. There is research currently underway to evaluate vignette performance in US academic vs. non-academic environments. Preliminary results indicate that vignette performance is consistent across settings and providers.[3]

WHAT IS A LESS DEVELOPED COUNTRY?

Given the review of problems in less developed countries, the reader can reasonably expect that the level of equipment in the facilities used in this study would vary significantly and thus possibly have an marked effect on quality. In the case of Macedonia, as presented in Chapter 3, although some of the problems present in less developed countries (poor health status of the population, heterogeneous access to health care, highly inconsistent health indicators among population segments, and inefficient management of health resources) are pervasive, it is still a country that devotes a high share of its GDP to health care. Moreover, at the time this study was conducted, the remnants of the centrally planned economy were still in place (Nordyke 2000), making it reasonable to expect a lower variability in levels of resources than a fully integrated market economy would have.

Thus it is important to take the results of this study into the context of a country that is less developed than industrialized western nations, but that shared some health indicators with developed nations at the time this study was conducted. It is also important to note that Macedonia is a country in transition, and that once market reforms are further implemented, a higher variation in health provision can be expected (Nordyke 2000).

---

[3] The author has been involved in undergoing research at five health facilities in California.

The type of care that will make a difference in a population's health status varies according to resource levels. The provision of basic health care has a relatively greater impact on low income or low education-or both-populations than does the procurement of more specialized services (Alderman 1996).

The contribution of this study centers on the methodology of quality assessment and analysis to formulate health policy. Each country needs to be analyzed according to their particular situation.

WHAT ARE WE REALLY MEASURING?

The analysis presented in this study is based on the explorations of aggregate data taken at a single point in time-or within a narrow interval. This poses the serious limitation of possibly not measuring exactly what the study should measure: the effects of quality on changes in health status.

By aggregating the data in a single point in time (i.e. a cross sectional study), it is not possible to assess the impact of clinical process on subsequent health. Clinical practice, particularly high quality, leads to a progression of physiologic changes that are manifested as better health.

From a public health point of view, however, aggregates are exactly what we need in order to gather information about possible changes in the health status of the population. By observing aggregate data on quality of care and health outcomes at a point in time, we may not be able identify the exact way in which immediate health problems are being addressed, but we have information on how the morbidity of chronic conditions is related to the provision of services. While policy cannot change health, it is intended to change behavior, particularly provider behavior, thereby improving the quality of the care and ultimately the health of the population.

IS THE CASE MIX CONTROLLED?

The strength of the analysis rests on the research design presented in Chapter 3, including "screened in" and "screened out" patients in the household survey controls for case mix, yet there may still be a possibility of bias when examining patients based on the exit procedure. Being able to use the households selected from the general population, would provide more assurance of effectively controlling for case mix. However, an effective link between medical facility and members of the population "not exited" does not exist, this is a serious study limitation.

DOES SELECTION BIAS EXIST?

A selection issue that is of concern and could potentially affect the relationship between process and outcome measures is the possibility that the clinics included in the study cater to different types of populations. Specifically, if there is a difference in income levels, there exist the possibility that clinics located in more affluent communities have populations with inherently better health and are also able to attract better physicians. For this purpose, a Kruskal-Wallis test of equality of income medians of communities served by the health facilities was conducted within each city. Only in Ohrid was there statistically significant evidence ($p < 0.05$) of a difference in median income between the different communities served by the health facilities. In the other three cities, there is no statistically significant evidence ($p < 0.005$) that median income varies by community. We can conclude that there are no significant economic differences between the different communities that will drive health outcomes and the process quality of practitioners.

Another selection issue is whether the healthcare facility users are a self-selecting group that includes people that are especially concerned with their health and therefore more likely to possess the habits that contribute to better health. In order to explore this

possibility, we conducted a comparison of objective health measures between users and non-users of health facilities

Table 5.1

Comparison of Health Indicators for Healthcare Facility Users vs. Non-Users

		Blood Pressure	Blood Glucose Content	Lung Capacity
Ohrid	Users	81.76	111.40	237.78
	Non-Users	77.11	87.43	421.60
Prilep	Users	82.31	125.32	278.83
	Non-Users	80.97	98.46	439.18
Shtip	Users	83.94	114.43	277.71
	Non-Users	77.34	101.09	366.68
Veles	Users	84.24	130.45	279.31
	Non-Users	78.72	104.07	421.83
Total	Users	82.92	119.08	267.20
	Non-Users	78.65	98.91	410.67

All differences are statistically significant at $p<0.001$ with the exception of Blood Pressure in Veles

Table 5.1 shows how the health indicators for the users of the health facilities are consistently poorer than the ones for non-users. This corroborates what intuition would signal: people visiting health care facilities are sicker. There is no statistical evidence though to support the idea that the patients selected in the exit survey represent people who are more concerned with their health and thus present an endogeneity problem for analysis.

The survey design thus allows us to make the comparisons we've made and to make inferences based on the results obtained. There is no statistical evidence of bias in the sampling procedure.

RECOMMENDATIONS FOR FURTHER STUDY

This research has provided insights into the effects of process of care on health status and can serve as a base for further study into the formulation of policy aimed at improving the provision of health care to populations in less developed countries. For that purpose, further study based on the findings presented in this dissertation can include an exploration into the relationship between process and changes in health status.

A study that effectively links changes in health outcomes with process should include time-series data on the progression and management of health indicators. Such data, coupled with the geographically aggregated physician process score, can allow us to examine in more detail the effects of the physician knowledge on health variation over time on a population basis.

The disadvantage posed by the paucity of process data from less developed countries can be compensated with the relatively lower costs of conducting surveys that can be expected. The use of process data collection instruments like clinical vignettes allows researchers to inquire into the levels of process of care even in the absence of patients' records or medical charts without recurring to more expensive and complicated methods like standardized patients.

The findings presented in this research provide a framework for the conduction of diagnostic studies of the provision of health care in a less developed country that lacks the information systems capable of providing process information. Further explorations into the workings of process can be undertaken with the present research as a basis.

CONCLUSION

In spite of the limitations in the study, there is validity to the empirical evidence provided on the variability and properties of the process scores. All doctors were examined using the same tools, and there is no reason to believe *a posteriori* that the clinical vignette scores are biased in any particular way.

The analysis of process scores in Macedonia present evidence that the quality of care variations exist across conditions, domains of care, and municipalities. This provides support for a policy intervention oriented towards improving their health care system.

The empirical evidence provided in this dissertation about the effect of a process measure on health status opens up the possibility for considering medical training and retraining as a higher priority than structural interventions in the efforts to improve a health care system.

The debate regarding the right intervention tools for the improvement of a health care system remains open however, and like any other system it is not realistic to consider policy levers on isolated terms.

Conceiving a good measure of quality of care in a health system is a requirement for sound policy formulation. Unfortunately, measuring quality is methodologically challenging. As a consequence, studies have tended to focus on the "structure" dimension of the quality of care (Donabedian, 1980), which is usually easier to quantify. For example, among World Bank projects funded from 1990-1993 that addressed quality of care issues in the developing world, quality was largely defined along the structure dimension (e.g., unreliable supply of essential drugs, disrepair of infrastructure and equipment). Such studies only marginally address or completely ignore the "process" dimension of

quality of care-e.g., inefficient outreach and patient referral systems, irrational use of drugs, high staff turnover, limited accessibility of family planning services, lack of coordination between levels of care (De Geyndt 1995). This is an important limitation of current policy-making, since many quality-of-care-related problems result from improper process of care (Bhat 1999, Brook 2000). It follows that many of the policy interventions to improve quality in a health system that need to be analyzed should address the process of care.

The stated research purpose of this study has been accomplished. We have presented empirical evidence of the effects of structural measures on process of care, and of the effects of medical knowledge on the health status of a population at the aggregate level.

Physician's characteristics played a greater role in explaining the process of care than equipment measures. For the outpatient conditions selected in this study, there is evidence to suggest that who the physicians are is a greater indicator of the level of care they will provide, than where they practice.

The process of care, in turn, is related to health status in the population, the ultimate mark of any health policy intervention. It is thus advisable that in seeking to improve a health system's performance, policymakers should try to explore interventions that target the process of care.

APPENDIX 1

BACKGROUND ON PROBLEMS OF LESS DEVELOPED COUNTRIES

The basic premise of analysis of development in this paper is that the process of economic growth is centered around a brief time interval where the economic organization of a society and its scale of values promote a set of preconditions that transform it such that economic growth is more or less automatic (Rostow 1956). The set of preconditons for economic growth include indicators of welfare and income distribution. As a brief overview of the development level and low standards of living existing in less developed countries, we have chosen three indicators of welfare that are not just related to the society's current situation, but also to its future prosperity. Social equality, consumption levels, and labor productivity are all signs of the structural situation of a society (Robalino 2000). They are also indicators of how likely it is that economic woes can be surmountable. The greater the social inequality, lower consumption levels, and lower labor productivity, the less likely a country is to overcome the problem of lower living standards.

In order to illustrate these indicators, I have gathered information on income distribution, consumption per capita, and labor activity rates for selected regions around the world.

The inequality in the distribution of income is illustrated by the share of either the income or the consumption accruing to a segment (percentage) of the population ranked by income distribution or consumption levels. All other things being equal, consumption tends to be a better indicator of welfare, especially in less developed countries. The Gini Coefficient[1] provides a suitable summary measure of the degree of distribution inequality.

---

[1] The Gini Coefficient measures the deviation from a perfectly equal distribution of income distribution among households within an economy. A Gini Coefficient of zero represents perfect equality, while an index of 100 means perfect inequality.

As an illustration of how the Gini coefficient represents distribution of income in a society, Table A1.1 presents the figures for the Gini coefficient and income distribution of quintiles ranked by income for selected countries.

Table A1.1

Gini Coefficient and Income Distribution in Selected Countries

Country	Gini Index	Lowest 20%	Second 20%	Third 20%	Fourth 20%	Highest 20%	Highest 10%
Australia	35.2	5.9	12.0	17.2	23.6	41.3	25.4
Austria	23.1	10.4	14.8	18.5	22.9	33.3	19.3
El Salvador	50.8	3.7	7.8	12.8	20.4	55.3	39.3
The Gambia	47.8	4.4	9.0	13.5	20.4	52.8	37.6
Philippines	46.2	5.4	8.8	13.2	20.3	52.3	36.6

Source: 2001 World Development Indicators, World Bank

The lower the Gini index for the countries on Table A1.1, the lower the percentage of income distributed among the people in the highest 10% of the income distribution. The most marked differences are observed in the highest quintile or the highest tenth, these figures provide a representation of the economic and social equality in a given country.

It can be argued as well, that the income distribution of the selected countries in this table is also an indicator of their level of economic development and their form of government and access to social services. For we can observe that those countries where income is less skewed in favor of the upper income tier are the countries with higher standards of living.

Figure A1.1 Income Distribution in Selected Regions
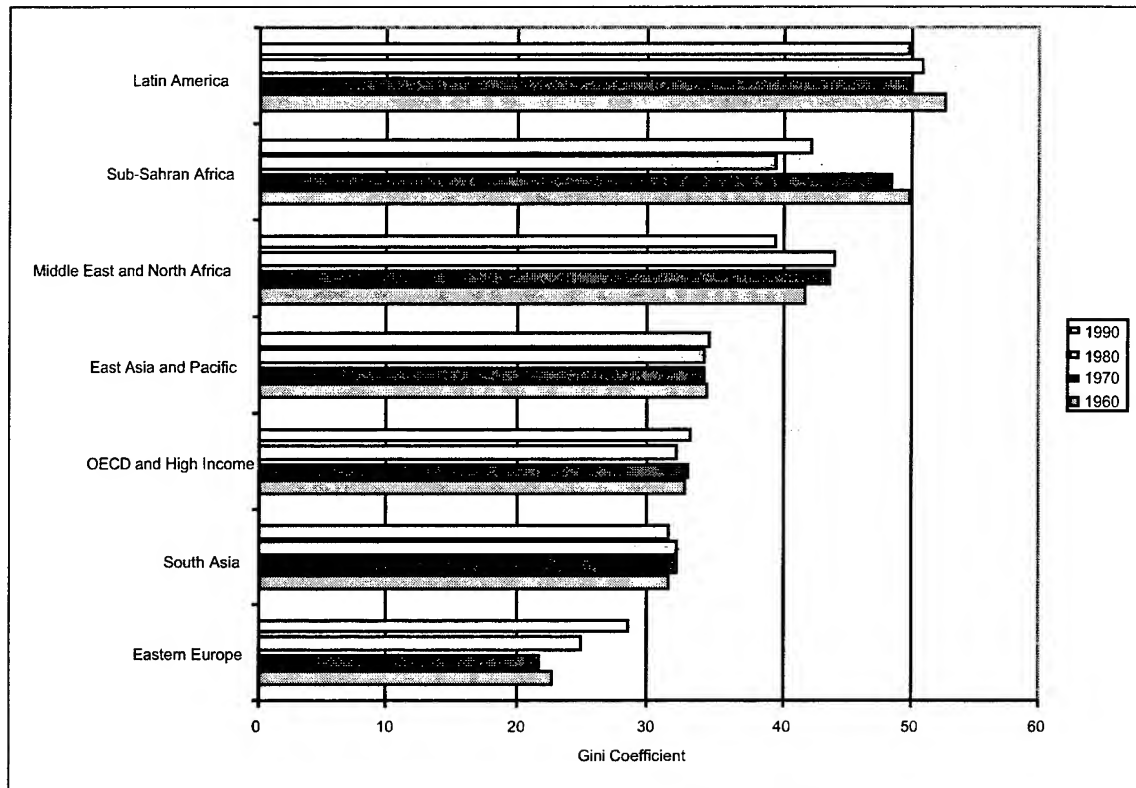
Source: 2001 World Development Indicators, World Bank

Figure A1.1 illustrates historical income distribution figures for different World regions. Less developed regions like Latin America and Sub-Saharan Africa present the most unequal income distributions. In these cases income is concentrated in few hands and therefore the inequality situation is perpetuated as lack of resources and opportunities prevent others from accessing the avenues for prosperity.

In examining Figure A1.1, a cautionary observation is that in regions that show historical fairness of distribution, even with respect to high income countries—like South Asia and Eastern Europe—a series of market-oriented reforms have been instituted in the past ten years. Historically, due to economic isolation due to cultural or political reasons, these countries' economies did not have the equity problems present in capitalist societies. As these societies open up and become part of the international economic markets, the implemented market reforms will on the average tend to create greater income disparities, especially as the mechanisms for the

functioning of markets are either not fully implemented or not adequately understood.

Taking the problems of income and consumption distribution into account, the figures for consumption are more sobering since the consumption attributable to countries where distribution is highly unequal overestimates the true consumption of the lowest ranks in the economic ladder. In those countries where income and consumption tend to be highly unequal, the people in the lowest segment may be receiving far less than their income per capita figures would indicate.

Table A1.2

Consumption per Capita in Selected Regions

		Low Income Countries	Middle Income Countries	High Income Countries
GNP per	WB Atlas (US$)[2]	420.00	1,980.00	26,440.00
capita / 1999	PPP (Inter. $)[3]	1,870.00	5,200.00	25,690.00

Source: 2001 World Development Indicators, World Bank

The consumption per capita in low income countries is markedly different than that for the other income tiers. Nationals of poor countries can, on the average, be expected to consume the equivalent of less than 8% of what citizens in high income countries consume. In other words, the average yearly consumption of a resident of a poor country amounts to less than the average monthly consumption of a resident in a high income country.

Of special concern is the dynamic of the consumption gap. While the gaps between developed economies and less developed countries are currently high, they can only expected to grow if current trends of economic growth continue over the next few years. In the last thirty years, consumption gaps between regions have only widened as the tools needed to succeed in an economy ever more reliant on technology, continue to be outside the grasp of the people in the poorest countries.

---

[2] World Bank Atlas Method
[3] Purchasing Power Parity Method

Figure A1.2 Dynamics of Consumption per Capita in Selected Regions
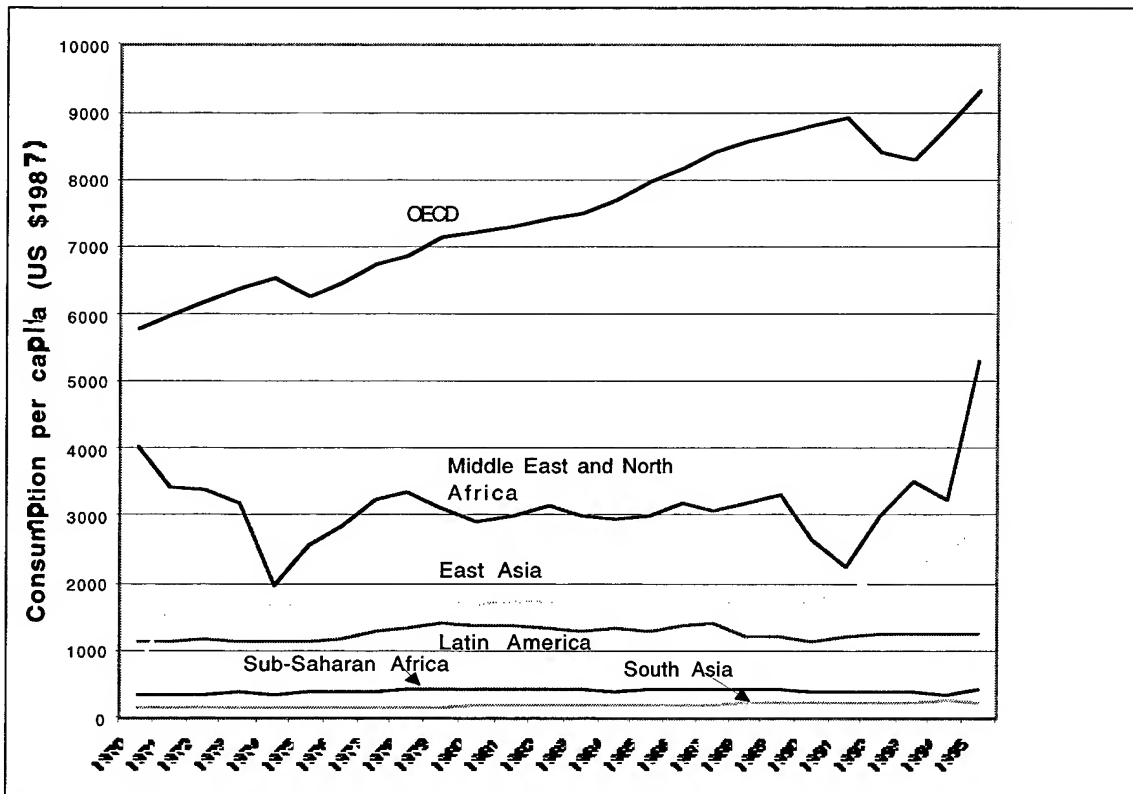
Source: Robalino 2000

Figure A1.2 presents the difference in consumption per capita in selected groups of countries. It can be appreciated from the figure that while in OECD countries consumption per capita has been rising steadily in the last thirty years, in less developed countries there has been a stagnation at levels that continue to increase the gap between the rich and the poor countries.

Countries in Sub-Saharan Africa and South Asia, with average consumption per capita below 1987 US$500.00 would need to steadily grow at rates above 20% per year for the next 20 years in order to achieve the current levels of consumption of OECD countries of around 1987 US$10,000.00. Conversely, for them to achieve those consumption per capita levels based on growth rates of 2% per year, it would take them close to two centuries.

In addition to the distribution inequalities within countries already mentioned, we also have to be cautious about the income inequalities within countries in the selected regions. Even though the grouping of countries has

been such that it presents somewhat similar figures, countries like Singapore and Hong Kong in East Asia, and Kuwait and U.A.E. in the Middle East/North Africa region, account for much of the increase in consumption per capita in their respective groups.

There are different reasons why the lower standards of living exist, and the socio-historical causes are not discussed in this paper, but an elucidation into the factors that directly affect the standards of living and the growth potential of less developed countries is provided.

As mentioned above, health and education are of importance, and one of the manifestations of how health and education interact to create welfare is in the productivity of labor.

Table A1.3

Unemployment and Labor Productivity in Selected Countries

	Unemployment 1996-98 % of total labor force	% of total unemployment by level of educational attainment 1996-98			Value Added per worker in Manufacturing 1995-99
		Primary	Secondary	Tertiary	
Bangladesh	2.5	47.4	28.4	9.9	1,711
Bolivia	4.2	24.1	42.3	29.0	26,282
Honduras	3.9	63.2	22.4	5.8	7,427
Ireland	7.8	64.5	23.9	10.9	86,036
Romania	6.3	21.3	72.1	5.9	3,482

Source: 2001 World Development Indicators, World Bank

Table A1.3 presents data from five selected countries at different stages in their development. The first column is their corresponding unemployment figures. Although generally considered an objective indicator of economic activity, cross-national comparison of unemployment levels by themselves can be a misleading indicator, since an initial observation of the figures may lead to conclude that the economic conditions in Bangladesh, Bolivia, and Honduras, are better than those in Ireland and Romania.

The misleading feature of unemployment by itself is that in countries without a developed government social network, many people eligible to be part of the workforce, when confronted with the prospect of unemployment, do opt

for subemployment, or employment in the informal sector, or just completely give up looking for a job. After all, the definition of unemployment only takes into account those actively seeking work. Therefore, the higher figures in unemployment for Ireland and Romania may be a reflection of a more developed public social network, and the fact that people can more reasonably expect to find a job in adequate time.

The composition of the educational segments of the unemployed gives an indication of the prospects for economic growth. As presented in Table 1.3, a high portion of the unemployed in countries like Bolivia and Romania are in the upper two educational tiers, this means that they can still improve productivity by incorporating higher skilled people to the labor force. Whereas countries like Bangladesh and Honduras don't have much latitude as far as employing more of their better educated workers since the portion of the unemployed who belong to the highest educational tier is relatively low.

In Ireland, the high portion of the unemployed represented by the people in the lower education tier indicates that jobs in their economy are given to the better educated. The portion corresponding to the better educated tier is higher than the other countries—except Bolivia—but not exceptionally high; this can be explained by frictional unemployment. The better educated find it easier to find good paying jobs so they are more likely to be between jobs.

Finally, value added per worker in manufacturing is a measure that does not need to be qualified to be compared, the higher the value added, the more productive the workforce. The caveat in using this figure is that it only applies to manufacturing activity, and a country that does not have a developed manufacturing industry but has other sources of income (agriculture, minerals) may appear to be at an economic disadvantage when it is not necessarily the case.

For comparative purposes among regions then, we will present the labor force activity rate. This is defined as the total labor force expressed as a percentage of the working-age population. This can provide us with a better illustration of the different rates at which people work in different regions. The comparison has been restricted to male participation in the workforce,

since female participation in the workforce cannot be expected to be homogeneous across regions due to cultural barriers which limit options for females in some countries (Peabody 1999a).

Thus, accounting for male participation only, the pattern of labor force activity in less developed countries appears lower than in developed nations.

Figure A1.3

Male Labor Force Activity in Selected Regions

Male Labor Force Activity Rate
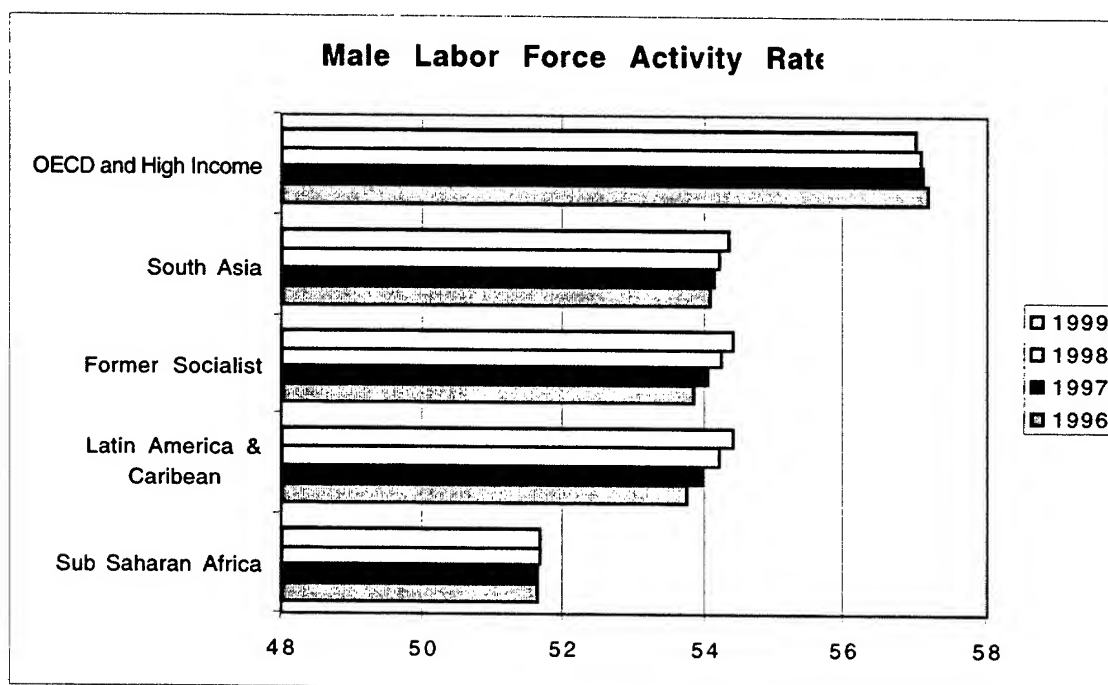
Source: 2001 World Development Indicators, World Bank

The figures for male labor force activity rate are consistently lower in less developed countries than in developed nations. This can only contribute to further deterioration of living conditions as less is produced and national resources—in the form of manpower—go unused.

Appendix 2

Clinical Vignette for Complex Low Back Pain:

- English Version

- Macedonian Version

- Scoring Criteria

Lb/pb

CLINICAL VIGNETTE - PRETEST SURVEY

Physician ID # [insert code label here]   Date: _____

SITE: Ohrid   Prilep   Veles   Shtip   (Please circle):

TYPE OF FACILITY (where you work most of the time)

    Medical Center   Health Center   Health Station   (Please circle):

Start Time: _____

End Time: _____

Specialty that best describes your practice: (Please check)

¡ General/Family Practice   ¡ Social Medicine

¡ Internal Medicine   ¡ Surgery

¡ Pediatrics   ¡ Surgical Specialist (e.g., ENT)

¡ Medical Pediatrics   ¡ Gynecologist/Obstetrician

¡ Specialty (e.g., cardiology)   ¡ Preventive Medicine

¡ Labor Medicine   ¡ Other: _____

COMMENTS

INSTRUCTIONS:

YOU HAVE BEEN GIVEN A CLINICAL VIGNETTE THAT TAKES APPROXIMATELY 10 MINUTES TO COMPLETE. IT IS SEPARATED INTO PARTS: ONCE YOU HAVE READ THE CLINICAL MATERIAL ON THE LEFT HAND PAGE, COMPLETE THE QUESTIONS ON THE FACING (RIGHT HAND) PAGE.

ONCE YOU HAVE ANSWERED THE QUESTIONS, TEAR THE PAGE OUT ALONG THE LEFT BINDING AND PLACE IT INTO THE MANILA ENVELOPE YOU HAVE BEEN GIVEN. IT IS IMPORTANT THAT YOU PLACE YOUR ANSWERS INTO THE ENVELOPE BEFORE GOING ON TO THE NEXT PART OF THE VIGNETTE.

PLEASE CONTINUE UNTIL ALL OF THE PAGES HAVE BEEN COMPLETED AND PLACED IN THE ENVELOPE. AS SOON AS YOU COMPLETE THIS VIGNETTE YOU MAY GO ON TO THE NEXT CLINICAL CASE IF YOU HAVE BEEN GIVEN MORE THAN ONE.

THANK YOU FOR YOUR PARTICIPATION IN THIS EVALUATION.

Lb/pb

READ THE VIGNETTE BELOW. ONCE YOU HAVE READ IT, PLEASE ANSWER THE QUESTIONS ON THE OPPOSITE PAGE.

A 55-year old man comes to the clinic after having developed back pain 2 weeks ago. He reports no inciting incident such as heavy lifting and his only activity had been gardening 2 days prior to onset of pain. The pain is described as dull, constant and 4 increasing to 6 out of 10 on a scale of 10 in the past 2 weeks, but it does awaken him at night. Rest is not helpful and the pain is worse with movement.

●●

PLEASE READ ALL FOUR QUESTIONS, THEN ANSWER UNDER THE MOST SUITABLE HEADING BELOW. DO NOT MOVE AHEAD TO THE NEXT PART OF THE VIGNETTE UNTIL ALL QUESTIONS HAVE BEEN ANSWERED ON THIS PAGE.

What are the 5 to 10 most important questions in the history you want to ask this patient about his back pain? (Please list.)

What are the 2 to 3 questions you want to know about his medications? (Please list.)

What are the 2 to 4 questions you want to ask about his past medical history? (Please list.)

What are the 3 to 7 most important questions you want to ask about his social history? (Please list.)

IMPORTANT: TEAR OUT ALONG LEFT BINDING. PLACE ANSWER SHEET IN MANILA ENVELOPE BEFORE CONTINUING.

Lb/pb

**PLEASE TURN PAGE ONCE YOU HAVE PLACED THE
ANSWER SHEET IN THE MANILA ENVELOPE.**

Lb/pb

CONTINUE READING THE VIGNETTE BELOW. ONCE YOU HAVE READ IT, PLEASE ANSWER THE QUESTIONS ON THE OPPOSITE PAGE

A 55-year old man comes to the clinic after having developed back pain 2 weeks ago. He reports no inciting incident such as heavy lifting and his only activity had been gardening 2 days prior to onset of pain. The pain is described as dull, constant and 4 increasing to 6 out of 10 on a scale of 10 in the past 2 weeks, but it does awaken him at night. Rest is not helpful and the pain is worse with movement.
●●●

The patient has had the pain for 2 weeks but denies any neurological symptoms or signs. He has had some relief of his pain with Tylenol or Advil. His medical history, however, is notable for localized prostate cancer 2 years ago confirmed by biopsy. The workup for metastatic disease was negative and the patient received RT for 6 weeks. His follow ups since then are every 6 months, and include a rectal exam and PSA which have been normal each time.

Twenty years ago he experienced a similar type of back pain that he saw the doctor for, received 'some medications' and the symptoms gradually resolved. He is otherwise well and does not have any other medical conditions including weight loss, fever, or IV drug use.

He works as a car mechanic and hasn't had to (and can't afford to) miss any work. He smokes 1/2 to 1 pack per day, consumes alcohol moderately, eats a "regular" diet and does not exercise other than his job.

His physical examination shows a temperature of 37° C, a blood pressure of 130/85, a pulse of 82. He walks stiffly and sits in the chair/gets up on the table gingerly.
●●●

**PLEASE ANSWER THE FOLLOWING QUESTIONS. DO NOT MOVE AHEAD TO THE NEXT
PART OF THE VIGNETTE UNTIL ALL QUESTIONS HAVE BEEN ANSWERED ON THIS PAGE.**

*What are the 8 to 12 most important elements of the physical examination that need to be
performed on this patient? (Please be as specific as possible about the finding you are looking
for.)*

Lb/pb

**PLEASE TURN PAGE ONCE YOU HAVE PLACED THE
ANSWER SHEET IN THE MANILA ENVELOPE.**

Lb/pb

A 55-year old man comes to the clinic after having developed back pain 2 weeks ago. He reports no inciting incident such as heavy lifting and his only activity had been gardening 2 days prior to onset of pain. The pain is described as dull, constant and 4 increasing to 6 out of 10 on a scale of 10 in the past 2 weeks, but it does awaken him at night. Rest is not helpful and the pain is worse with movement.
•••

The patient has had the pain for 2 weeks but denies any neurological symptoms or signs. He has had some relief of his pain with Tylenol or Advil. His medical history, however, is notable for localized prostate cancer 2 years ago confirmed by biopsy. The workup for metastatic disease was negative and the patient received RT for 6 weeks. His follow ups since then are every 6 months, and include a rectal exam and PSA which have been normal each time.

Twenty years ago he experienced a similar type of back pain that he saw the doctor for, received 'some medications' and the symptoms gradually resolved. He is otherwise well and does not have any other medical conditions including weight loss, fever, or IV drug use.

He works as a car mechanic and hasn't had to (and can't afford to) miss any work. He smokes 1/2 to 1 pack per day, consumes alcohol moderately, eats a "regular" diet and does not exercise other than his job.

His physical examination shows a temperature of 37° C, a blood pressure of 130/85, a pulse of 82. He walks stiffly and sits in the chair/gets up on the table gingerly.
•••
The physical examination was unremarkable except for some midline tenderness in the midlumbar area and bilateral paraspinous muscle pain. Straight leg raising test did not elicit pain until legs raised to 60 degrees when patient experiences pain in the back of his upper thighs. The reflexes, strength and sensation were all normal.
•••

**PLEASE ANSWER THE FOLLOWING QUESTIONS. DO NOT MOVE AHEAD TO THE NEXT
PART OF THE VIGNETTE UNTIL ALL QUESTIONS HAVE BEEN ANSWERED ON THIS PAGE.**

At this point, what laboratory tests/imaging studies would you order?

**IMPORTANT: TEAR OUT ALONG LEFT BINDING. PLACE
ANSWER SHEET IN MANILA ENVELOPE BEFORE CONTINUING.**

Lb/pb

Lb/pb

**PLEASE TURN PAGE ONCE YOU HAVE PLACED THE
ANSWER SHEET IN THE MANILA ENVELOPE.**

Lb/pb

PLEASE CONTINUE READING THE VIGNETTE BELOW. ONCE YOU HAVE READ IT, PLEASE ANSWER THE QUESTIONS ON THE OPPOSITE PAGE.

A 55-year old man comes to the clinic after having developed back pain 2 weeks ago. He reports no inciting incident such as heavy lifting and his only activity had been gardening 2 days prior to onset of pain. The pain is described as dull, constant and 4 increasing to 6 out of 10 on a scale of 10 in the past 2 weeks, but it does awaken him at night. Rest is not helpful and the pain is worse with movement.

••

The patient has had the pain for 2 weeks but denies any neurological symptoms or signs. He has had some relief of his pain with Tylenol or Advil. His medical history, however, is notable for localized prostate cancer 2 years ago confirmed by biopsy. The workup for metastatic disease was negative and the patient received RT for 6 weeks. His follow ups since then are every 6 months, and include a rectal exam and PSA which have been normal each time.

Twenty years ago he experienced a similar type of back pain that he saw the doctor for, received 'some medications' and the symptoms gradually resolved. He is otherwise well and does not have any other medical conditions including weight loss, fever, or IV drug use.

He works as a car mechanic and hasn't had to (and can't afford to) miss any work. He smokes 1/2 to 1 pack per day, consumes alcohol moderately, eats a "regular" diet and does not exercise other than his job.

His physical examination shows a temperature of 37° C, a blood pressure of 130/85, a pulse of 82. He walks stiffly and sits in the chair/gets up on the table gingerly.

••

The physical examination was unremarkable except for some midline tenderness in the midlumbar area and bilateral paraspinous muscle pain. Straight leg raising test did not elicit pain until legs raised to 60 degrees when patient experiences pain in the back of his upper thighs. The reflexes, strength and sensation were all normal.

••

The laboratory and imaging studies are all within normal limits.

••

PLEASE ANSWER THE FOLLOWING QUESTIONS. DO NOT MOVE AHEAD TO THE NEXT VIGNETTE UNTIL ALL QUESTIONS HAVE BEEN ANSWERED ON THIS PAGE.

What would be your diagnosis?

And what would be your plan, including recommendations to the patient?

IMPORTANT: TEAR OUT ALONG LEFT BINDING. PLACE ANSWER SHEET IN MANILA ENVELOPE BEFORE CONTINUING.

Lb/pb

KLINI^KA VIWETA- anketa za pretestot

Broj Na Slu~aj:

_____   **Vreme na po~etok**

_____   **Vreme na kraj**

KOMENTARI

INSTRUKCII:

DOSTAVENA VI E KLINI^KA VIWETA ZA KOJA E POTREBNO OKOLU 10 MINUTI DA SE POPOLNI. PODELENA E NA NEKOLKU DELA: OTKAKO JE GO PRO^ITATE KLINI^KIOT MATERIJAL NA STRANICATA OD LEVATA STRANA, ODGOVORETE GI PRA[AWATA NA STRANICATA OD DESNATA STRANA.

OTKAKO JE GI ODGOVORITE PRA[AWATA, SKINETE JA STRANICATA PO LEVATA MARGINA I STAVETE JA VO PLIKOT [TO VI E DADEN. VA@NO E DA GI STAVITE VA[ITE ODGOVORI VO PLIKOT <u>PRED</u> DA PREMINETE NA SLEDNIOT DEL OD VIWETATA.

VE MOLIME, PRODOL@ETE DODEKA NE GI POPOLNITE SITE STRANICI I ISTITE NE GI STAVITE VO PLIKOT. VEDNA[[TOM JE JA POPOLNITE OVAA VIWETA, MO@ETE DA PREMINETE NA SLEDNIOT KLINI^KI SLU^AJ, DOKOLKU E DADEN VO VIWETATA.

VI BLAGODARIME ZA VA[ETO U^ESTVO VO OVAA EVALUACIJA.

16   Lb/pb

PRO^ITAJTE JA VIWETATA PODOLU. [TOM]E JA PRO^ITATE, VE MOLIME DA GI ODGOVORITE PRA[AWATA OD SPROTIVNATA STRANICA.

55 godi{en ~ovek doa|a vo klinikata otkako dobil bolki vo grbot pred dve nedeli. Toj vi ka`uva deka nemal nikakov provocira~ki incident kako {to e digawe te`ok tovar i negovata edinstvena aktivnost dva dena, pred da se javat bolkite, bila rabota vo gradina. Bolkata e opi{ana kako tapa, postojana, i na skala do 10, taa bi bila 4 {to odi kon 6 vo poslednite dve nedeli i taa go budi no}e. Odmoraweto ne pomaga i bolkata se vlo{uva pri dvi`ewe.

•••

VE MOLAM PRO^ITAJTE GI SITE ^ETIRI PRA[AWA, A POTOA ODGOVORETE POD NAJSOODVETNOTO ZAGLAVJE OD DOLUNAVEDENITE. NE PREMINUVAJTE NA SLEDNIOT DEL OD VIWETATA SE DODEKA NE GI ODGOVORITE SITE PRA[AWA OD OVAA STRANICA.

Koi se najva`nite 5 do 10 pra[awa od anamnezata {to bi sakale da mu gi postavite na ovoj pacient vo vrska so negovata bolka vo grbot? (Ve molime, navedete)

Koi se najva`nite 2 do 3 pra[awa {to sakate da gi doznaete vo vrska so negovite lekarstva? (Ve molime, navedete.)

Koi se najva`nite 2 do 4 pra[awa {to sakate da gi doznaete vo vrska so negovite minati zaboluvawa? (Ve molime, navedete.)

Koi se najva`nite 3 do 7 pra[awa {to sakate da gi doznaete za negoviot stil na `iveewe? (Ve molam, navedete.)

VA@NO: SKINETE PO LEVATA MARGINA.STAVETE JA STRANICATA SO ODGOVORITE VO MANILSKIOT PLIK PRED DA PRODOL@ITE.

Lb/pb

VE MOLIME, SVRTETE JA OVAA STRANICA OTKAKO VEJE STE JA
STAVILE STRANICATA SO ODGOVORITE VO PLIKOT.

55 godi{en ~ovek doa|a vo ambulantata otkako dobil bolki vo grbot pred dve nedeli. Toj vi ka`uva deka nemal nikakov provocira~ki incident kako {to e digawe te`ok tovar i negovata edinstvena aktivnost dva dena, pred da se javat bolkite, bila rabota vo gradina. Bolkata e opi{ana kako tapa, postojana, i na skala do 10, taa bi bila 4 {to odi kon 6 vo poslednite dve nedeli i taa go budi no}e. Odmoraweto ne pomaga i bolkata se vlo{uva pri dvi`ewe.

●●●

Pacientot ima bolki ve}e dve nedeli, no negira sekakvi neurolo{ki simptomi ili znaci. Paracetamolot i aspirinot mu ja olesnuvaat malku bolkata. Me|utoa, negovata anamneza uka`uva deka imal lokaliziran rak na prostata pred dve godini, potvrden so biopsija. Ispituvawata za metastazite bea negativni i pacientot dobil terapija so radijacija vo traewe od {est nedeli. Ottoga{ negovite kontrolni pregledi vklu~uvaat rektalen pregled i ispituvawe na specifi~en antigen na prostata se vr{at na sekoi 6 meseci i bile normalni pri sekoja kontrola.

Pred dvaeset godini imal sli~na bolka vo grbot i toga{ se javil na lekar i primil "nekakvi lekarstva" i simptomite postepeno is~eznale. Inaku, toj e zdrav i nema drugi hroni~ni oboluvawa vklu~uvaj}i gubewe na te`ina, treska, ni intervenozno koristewe na lekarstva.

Toj raboti kako avtomehani~ar i ne moral (i ne mo`e da si dozvoli finansiski) da otsustvuva od rabota. Pu{i polovina do edna kutija cigari dnevno, konsumira alkohol umerno, jade "normalna" hrana i ne pravi fizi~ki ve`bi osven ona {to mu e vrzano za rabotata.

Negoviot lekarski pregled poka`a deka ima temeratura od 37 stepeni, krven pritisok od 130/85, puls od 82. Toj se dvi`i zdrveno i sedi i stanuva od stol vnimatelno.

●●●

VE MOLIME DA GI ODGOVORITE SLEDNITE PRA[AWA. NE PREMINUVAJTE NA SLEDNATA STRANICA NA VIWETATA DODEKA NE GI ODGOVORITE PRA[AWATA NA OVAA STRANICA.

Koi se najva`nite 8 do 12 elementi na lekarskiot fizikalen  pregled {to treba da mu se napravat na pacientot? (Ve molime, bidete kolku {to e mo`no poprecizni vo baraweto na naodite)

VA@NO: SKINETE JA STRANICATA PO LEVATA MARGINA. STAVETE JA STRANICATA SO ODGOVORITE VO PLIKOT PRED DA PRODOL@ITE.

Lb/pb

Lb/pb

VE MOLIME, SVRTETE JA OVAA STRANICA OTKAKO VE]E STE JA STAVILE STRANICATA SO ODGOVORITE VO PLIKOT.

Lb/pb

PRODOL@ETE DA JA ^ITATE VIWETATA PODOLU. OTKAKO]E JA PRO^ITATE, VE MOLIME DA GI ODGOVORITE PRA[AWATA NA SPROTIVNATA STRANICA.

55 godi{en ~ovek doa|a vo ambulantata otkako dobil bolki vo grbot pred dve nedeli. Toj vi ka`uva deka nemal nikakov provocira~ki incident kako {to e digawe te`ok tovar i negovata edinstvena aktivnost dva dena, pred da se javat bolkite, bila rabota vo gradina. Bolkata e opi{ana kako tapa, postojana, i na skala do 10, taa bi bila 4 {to odi kon 6 vo poslednite dve nedeli i taa go budi no}e. Odmoraweto ne pomaga i bolkata se vlo{uva pri dvi`ewe.

....

Pacientot ima bolki ve}e dve nedeli, no negira sekakvi neurolo{ki simptomi ili znaci. Paracetamolot i aspirinot mu ja olesnuvaat malku bolkata. Me|utoa, negovata anamneza uka`uva deka imal lokaliziran rak na prostata pred dve godini, potvrden so biopsija. Ispituvawata za metastazite bea negativni i pacientot dobil terapija so radijacija vo traewe od {est nedeli. Ottoga{ negovite kontrolni pregledi vklu~uvaat rektalen pregled i ispituvawe na specifi~en antigen na prostata, se vr{at na sekoi 6 meseci i bile normalni pri sekoja kontrola.

Pred dvaeset godini imal sli~na bolka vo grbot i toga{ se javil na lekar i primil "nekakvi lekarstva" i simptomite postepeno is~eznale. Inaku, toj e zdrav i nema drugi hroni~ni oboluvawa vklu~uvaj}i gubewe na te`ina, treska, ni intervenozno koristewe na lekarstva.

Toj raboti kako avtomehani~ar i ne moral (i ne mo`e da si dozvoli finansiski) da otsustvuva od rabota. Pu{i polovina do edna kutija cigari dnevno, konsumira alkohol umerno, jade "normalna" hrana i ne pravi fizi~ki ve`bi osven ona {to mu e vrzano za rabotata.

Negoviot lekarski pregled poka`a deka ima temeratura od 37 stepeni, krven pritisok od 130/85, puls od 82. Toj se dvi`i zdrveno i sedi i stanuva od stol vnimatelno.

....

Rezultatite od lekarskiot pregled bea normalni, osven {to ima{e slaba ~uvstvitelnost vo sredi{niot lumbalen del i bilateralno se javuva paraspinozna bolka vo muskulite. Testot so ispraveni noze ne isprovocira bolka se dodeka ne gi digna nozete na 60 stepeni i toga{ pacientot po~uvstvuva bolka pozadi vo gorniot del na natkolenicata. Refleksite, ja~inata i senzitivnosta bea normalni.

....

Lb/pb

VE MOLIME, ODGOVORETE GI SLEDNITE PRA[AWA. NE MINUVAJTE NA SLEDNATA STRANICA NA VIWETATA DODEKA NE GI ODGOVORITE PRA[AWATA OD OVAA STRANICA.

Vo ovoj stadium, koi laboratoriski testovi/ispituvawa (rentgen, EHO i dr.) bi gi napravile?

VA@NO: SKINETE JA STRANICATA PO LEVATA MARGINA. STAVETE JA STRANICATA SO ODGOVORITE VO PLIKOT PRED DA PRODOL@ITE.

VE MOLIME SVRTETE JA OVAA STRANICA OTKAKO VEЈЕ STE JA
STAVILE STRANICATA SO ODGOVORITE VO PLIKOT.

Lb/pb

PRODOL@ETE DA JA ^ITATE VIWETATA PODOLU. OTKAKO JE JA PRO^ITATE, VE MOLIME DA GI ODGOVORITE PRA[AWATA NA SPROTIVNATA STRANICA.

55 godi{en ~ovek doa | a vo ambulantata otkako dobil bolki vo grbot pred dve nedeli. Toj vi ka`uva deka nemal nikakov provocira~ki incident kako {to e digawe te`ok tovar i negovata edinstvena aktivnost dva dena, pred da se javat bolkite, bila gradinarstvoto. Bolkata e opi{ana kako tapa, postojana, i na skala do 10, taa bi bila 4 {to odi kon 6 vo poslednite dve nedeli i taa go budi no}e. Odmoraweto ne pomaga i bolkata se vlo{uva pri dvi`ewe.

••

Pacientot ima bolki ve}e dve nedeli, no negira sekakvi neurolo{ki simptomi ili znaci. Paracetamolot i aspirinot mu ja olesnuvaat malku bolkata. Me | utoa, negovata anamneza uka`uva deka imal lokaliziran rak na prostata pred dve godini, potvrden so biopsija. Ispituvawata za metastazite bea negativni i pacientot dobil terapija so radijacija vo traewe od {est nedeli. Ottoga{ negovite kontrolni pregledi vklu~uvaat rektalen pregled i ispituvawe na specifi~en antigen na prostata se vr{at na sekoi 6 meseci i bile normalni pri sekoja kontrola.

Pred dvaeset godini imal sli~na bolka vo grbot i toga{ se javil na lekar i primil "nekakvi lekarstva" i simptomite postepeno is~eznale. Inaku, toj e zdrav i nema drugi hroni~ni oboluvawa vklu~uvaj}i gubewe na te`ina, treska, ni intervenozno koristewe na lekarstva.

Toj raboti kako avtomehani~ar i ne moral (i ne mo`e da si dozvoli finansiski) da otsustvuva od rabota. Pu{i polovina do edna kutija cigari dnevno, konsumira alkohol umerno, jade "normalna" hrana i ne pravi fizi~ki ve`bi osven ona {to mu e vrzano za rabotata.

Negoviot lekarski pregled poka`a deka ima temeratura od 37 stepeni, krven pritisok od 130/85, puls od 82. Toj se dvi`i zdrveno i sedi i stanuva od stol vnimatelno.

••

Rezultatite od lekarskiot pregled bea normalni, osven {to ima}e slaba ~uvstvitelnost vo sredi{niot lumbalen del i bilateralno se javuva paraspinozna bolka vo muskulite. Testot so ispraveni noze ne isprovocira bolka se dodeka ne gi digna nozete na 60 stepeni i toga{ pacientot po~uvstvuva bolka pozadi vo gorniot del na natkolenicata. Refleksite, ja~inata i senzitivnosta bea normalni.

••

Laboratoriskite rezultati i ispituvawa se vo granicite na normalata.

••

Lb/pb

VE MOLIME, ODGOVORETE GI SLEDNIVE PRA[AWA. NE ODETE NA SLEDNATA
VIWETA SE DODEKA NE GI ODGOVORITE SITE PRA[AWA NA OVAA STRANICA.

Koja bi bila va{ata dijagnoza?

I {to bi opfatil va{iot plan, vklu~uvaj}i gi i preporakite za pacientot?

Lb/pb

CLINICAL VIGNETTE

SCORING SHEET
LBP-2 (L/bpb)

(September 2, 1997)

1. Date Administered: ___ ___ / ___ ___ / ___ ___   1-6/

2. Physician: ___ ___   7-8/

3. Site: ___   9/

4. Case: ___   10/

Physician ID # ____ ____

LBP–2 Vignette Scoring Sheet

What are the most important questions in the history you want to ask this patient about his back pain?

_____ ¡ yes ¡ no   ¡

yes   ¡ no   11/
1(1). how long have you had the back pain *(2 weeks)*.

_____ ¡ yes ¡ no   ¡

yes   ¡ no   12/
2(2). any radiation of the pain *(none)*.

_____ ¡ yes ¡ no   ¡

yes   ¡ no   13/
3(3). any previous history of back pain *(20 years ago)*.

_____ ¡ yes ¡ no   ¡

yes   ¡ no   14/
4(4). any history of cancer, trauma, weight loss, fever, IV drug use, or prolonged
          steroid use *(prostate cancer 2 years ago)*.

_____ ¡ yes ¡ no   ¡

yes   ¡ no   15/
5(5). any weakness, numbness or tingling *(none)*.

_____ ¡ yes ¡ no   ¡

yes   ¡ no   16/
6(6). any changes in bladder or bowel functions *(none)*.

_____ ¡ yes ¡ no   ¡

yes   ¡ no   17/
7(9). any loss of work due to the pain *(none, can't afford to miss)*.

_____

_____

_____

Physician ID # ____ ____

What are the most important questions you want to ask about his past medical history?

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   18/
8(8). any other medical conditions (diabetes, hypertension, high cholesterol, etc.) *(none)*.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   19/
9(16). previous preventive health care *(PSA, rectal, etc.)*.

_____

_____

What are the most important questions you want to ask about his medications?

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   20/
10(7) any medications *(Tylenol and Advil, for current pain only)*.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   21/
11(11). drug allergies *("none")*.

_____

_____

Lb/pb

What are the most important questions you want to ask about him as a person (his social history)?

_____   ¡ yes ¡ no   ¡

yes   ¡ no   22/

12(10). type of work *(auto mechanic)*.

_____   ¡ yes ¡ no   ¡

yes   ¡ no   23/

13(12). smoking history.

_____   ¡ yes ¡ no   ¡

yes   ¡ no   24/

14(13). alcohol use.

_____   ¡ yes ¡ no   ¡

yes   ¡ no   25/

15(14). diet.

_____   ¡ yes ¡ no   ¡

yes   ¡ no   26/

16(15). amount of exercise *("none")*.

_____

_____

What are the most important elements of the physical examination that need to be performed on this patient?

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   27/
17(17). examine eyes, ears, nose, and/or throat.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   28/
18(18). feel up and down spine.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   29/
19(19). press along muscles on either side of spine.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   30/
20(20). ask to bend forward at the waist.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   31/
21(21). ask to bend to both sides at the waist.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   32/
22(22). do straight leg raises on both sides while lying on back.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   33/
23(23). check for reflexes at the knees on both legs.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   34/
24(24). check for reflexes at the ankles on both legs.

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   35/
25(25). ask to push down or pull up against hand with foot (or great toe).

_____   ¡ yes   ¡ no   ¡
yes   ¡ no   36/
26(26). check for sensation on both legs.

Lb/pb

_____ j yes j no   j

yes   j no   37/

27(27). ask to do a rectal exam (patient will refuse).

_____

_____

Physician ID # ____ ____

At this point, what laboratory tests/imaging studies would you order?

_____ ¡ yes ¡ no   ¡
yes   ¡ no   38/
28(X). Lumbar spine series.

_____ ¡ yes ¡ no   ¡
yes   ¡ no   39/
29(X). Bone scan.

_____ ¡ yes ¡ no   ¡
yes   ¡ no   40/
30(X). MRI.

_____ ¡ yes ¡ no   ¡
yes   ¡ no   41/
31(X). CT scan.

_____ ¡ yes ¡ no   ¡
yes   ¡ no   42/
32(X). PSA.

_____ ¡ yes ¡ no   ¡
yes   ¡ no   43/
33(X). Sed rate.

_____ ¡ yes ¡ no   ¡
yes   ¡ no   44/
34(X). Routine laboratory tests (CBC, LFTs, Chem, etc.).

_____

_____

_____

Lb/pb

Physician ID # ____ ____

What would be your diagnosis?

_____ ¡ yes ¡ no   ¡

yes   ¡ no 45/

35(28). low back pain, with indications to rule out metastatic disease.

_____ ¡ yes ¡ no   ¡

yes   ¡ no 46/

36(x). discuss/mention the severity of back pain (mild, moderate, severe, etc.).

_____

_____

And what would be your plan, including recommendations to the patient?

_____ ¡ yes ¡ no   ¡

yes   ¡ no 47/

37(29). explain that more tests may be needed.

_____ ¡ yes ¡ no   ¡

yes   ¡ no 48/

38(31). discuss/demonstrate back exercises.

_____ ¡ yes ¡ no   ¡

yes   ¡ no 49/

39(32). discuss/demonstrate how to protect back when doing activities (i.e., lifting, etc.).

_____ ¡ yes ¡ no   ¡

yes   ¡ no 50/

40(33). recommend some type of aerobic activity (i.e., walking, running, etc.).

_____ ¡ yes ¡ no   ¡

yes   ¡ no 51/

41(34). recommend bedrest.

                         how long? _____   52/

_____ ¡ yes ¡ no   ¡

yes   ¡ no 53/

42(35). discuss the importance of stopping smoking.

Lb/pb

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   54/

43(36). recommend some type of follow-up.

<center>*when?* _____</center>

55

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   56/

44(X). prescribe acetaminophen.

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   57/

45(X). prescribe muscle relaxants..

Lb/pb

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   58/

46(X). prescribe narcotic analgesic.

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   59/

47(X). prescribe NSAID.

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   60/

48(X). suggest chiropractic referral.

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   61/

49(X). initiate physical therapy referral.

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   62/

50(X). initiate orthopedic consultation.

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   63/

51(X). initiate neurosurgery consultation.

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   64/

52(X). initiate PMR consultation.

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   65/

53(X). initiate rheumatology consultation.

_____   ¡ yes   ¡ no   ¡

yes   ¡ no   66/

54(X). perform or provide for preventive measures (influenza or DT vaccine,
         pneumovax, colon cancer screening, cholesterol, etc.).

_____

_____

Lb/pb

REFERENCES

Akin, John S., et al, "The Demand for Adult Outpatient Service in the Bicol Region of the Philippines," *Economic Development and Cultural Change*, Vol. 34, 1986, pp. 755-782.

Akin, John S., et al, "The Demand for Child Health Services in the Philippines," *Social Science & Medicine*, Vol. 13C, 1981, pp. 249-257.

Akin, John S., et al, *The Demand for Primary Health Services in the Third World*, New Jersey: Rowman & Allanheld, 1985.

Alderman, Harold and Victor Lavy, "Household Responses to Public Health Services: Cost and Quality Tradeoffs," *The World Bank Research Observer*, Vol. 11, No. 1, 1996, pp. 3-22.

Amemiya, Takeshi, "Qualitative Response Models: A Survey," *Journal of Economic Literature*, Vol. 19, No. 4, 1981, pp. 1483-1536.

Arrow, Keneth J, "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review*, Vol. 53, No. 5, 1963, pp. 941-973.

Badger, L.W., et al, "Stability of Standardized Patients' Performance in a Study of Clinical Decision Making," *Fam Med*, Vol. 27, 1995, pp. 126–131.

Batalden, Paul B., "Linking Outcomes Measurement to Continual Improvement: The Serial 'V' Way of Thinking About Improving Clinical Care," *Journal on Quality Improvement*, Vol. 20, No. 4, 1994, pp. 167-177.

Berman, Peter and Suomi Sakai, "The Productivity of Manpower and Supplies in Public Health Services in Java," in Anne Mills and Kenneth Lee, eds., *Health Economics Research in Developing Countries*, New York: Oxford University Press, 1993, pp. 314–334.

Bernard, Michael H., et al, "Patient Satisfaction in Developing Countries," *Social Science & Medicine*, Vol. 48, 1999, pp. 989-996.

Berwick, Donald M, "Controlling Variation in Health Care: A Consultation from Walter Shewhart," *Medical Care*, Vol. 29, No. 12, 1991, pp. 1212-1225.

Bhat, Ramesh, "Characteristics of Private Medical Practice in India: A Provider Perspective," *Health Policy and Planning*, Vol. 14, No. 1, 1999, pp. 26-37.

Bhat, Ramesh, "Regulation of the Private Health Sector in India," *International Journal of Health Planning and Management*, Vol. 11, 1996, pp. 253-274.

Bitran, R., *A Household Health Care Demand Study in Kisantu and Bokoro Regions of Zaire, Volume III*, Washington D.C.: World Bank, Resources for Child Health Project Working Paper, 1989.

Brook, Robert, and Kathleen Lohr, *Efficacy, Effectiveness, Variations, and Quality: Boundary-Crossing Research*, Santa Monica, CA: RAND, N-3368-HHS, 1991. Reprinted from *Medical Care*, Vol. 23, No. 5, pp. 710-722.

Brook, Robert, et al, "Defining and Measuring Quality of Care: a Perspective from US Researchers," *International Journal of Quality of Health Care*, Vol. 12, No. 4, 2000, pp. 281-295.

Brook, Robert, et al, "Geographic Variations in the Use of Services: Do they have any Clinical Significance?," *Health Affairs*, Vol. 3, No. 2, 1984, pp. 63-73.

Brook, Robert, et al, "Predicting the Appropriate Use of Carotid Endarterectomy, Upper Gastrointestinal Endoscopy, and Coronary Angiography," *NEJM*, Vol. 323, 1990, pp. 1173-1177.

Brook, Robert, et al, "Quality of Health Care, Part 2: Measuring Quality of Care," *NEJM*, Vol. 335, No. 13, 1996, pp. 966-970.

Chassin, Mark R., et al, "Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services? A Study of Three Procedures," *JAMA*, Vol. 258, No. 18, 1987, pp. 2533-2537.

Colliver, J.A., "Validation of Standardized-Patient Assessment: A Meaning for Clinical Competence," *Academic Medicine*, Vol. 70, 1995, pp. 1062-1064.

Colliver, J.A., et al, "Effects of the Examinee Gender Standardized-Patient Gender, and their Interaction on Standardized Patient's Ratings of Examinees Interpersonal and Communications Skills *Academic Medicine*, Vol. 68, 1993, pp. 153-157.

DaVanzo, J, and Paul Gertler, *Household Production of Health: A Microeconomic Perspective on Health Transitions*, Santa Monica, CA: RAND, N-3014-RC, 1990.

De Champlain, A.F, et al, "Standardized Patients' Accuracy in Recording Examinees' Behaviors Using Checklists," *Academic Medicine*, Vol. 72 (Suppl1), 1997, pp. S85-S87.

De Geynt, Willy, *Managing the Quality of Health Care in Developing Countries*, Washington DC: The World Bank, Technical Paper # 258, 1995.

Donabedian, Avedis, *The Definition of Quality and Approaches to its Assessment, Volume I*, Ann Arbor: Health Administration Press, 1980.

Dor, A., and J. Van der Gaag, *The Demand for Medical Care in Developing Countries: Quantity Rationing in Rural Côte d'Ivorie*, Washington DC: The World Bank, Living Standards Measurement Study Working Paper # 35, 1987.

Dosi, Giovanni, "Sources, Procedures, and Microeconomic Effects of Innovation," *The Journal of Economic Literature*, Volume 26, September 1988, pp. 1120-71.

Dresselhaus, Timothy, et al, "Measuring Compliance with Preventive Care Guidelines: Standardized Patients, Clinical Vignettes, and the Medical Record," Journal of General Internal Medicine, Vol. 15, No. 11, 2000, pp. 782–788.

Durán-Arenas, Luis, et al, "The Development of a Quality Information System: a Case Study of Mexico," *Health Policy and Planning*, Vol. 13, No. 4, 1998, pp.446-458.

Eddy, D.M., "Variation in Physician Practice: The Role of Uncertainty," *Health Affairs*, Vol. 3, No. 2, 1984, pp. 74-89.

Evans, Robert G., et al, eds., *Why Are Some People Healthy and Others Not?*, New York: Aldine de Gruyter, 1994.

Evidence-Based Medicine Working Group, "Evidence-Based Medicine," *JAMA*, Vol. 268, No. 17, 1992, pp.2420–2425.

Ford, HL, et al, "Health status and quality of life of people with multiple sclerosis," *Disabil Rehabil*, Vol. 23, No. 15, 2001, pp. 516-21.

Foster, Susan, "Economic Aspects of the Production and Use of Pharmaceuticals: Evidence and Gaps in Research," in Anne Mills and Kenneth Lee, eds., *Health Economics Research in Developing Countries*, New York: Oxford University Press, 1993, pp. 288-313.

Gani, Ascobat, "Improving Quality in Public Sector Hospitals in Indonesia," *International Journal of Health Planning and Management*, Vol. 11, 1996, pp. 274-296.

Gertler, Paul, and Jacques van der Gaag, *The Willingness to Pay for Medical Care, Baltimore: Evidence from Two Developing Countries*, MD: The Johns Hopkins University Press, 1990.

Ghana Statistical Service (GSS) and Macro International Inc. (MI), *Ghana Demographic and Health Survey 1998*, Calverton, Maryland: GSS and MI, 1999.

Gill, Richard D., et al, "Large Sample Theory of Empirical Distributions in Biased Sampling Models," *The Annals of Statistics*, Vol. 16, No. 3, 1988, pp.1069–1112.

Gillis, Malcom, et al, *Economics of Development*, New York: W. W. Norton & Company, 1992.

Glassman, Peter A., et al, "Using Standardized Patients to Measure Quality: Evidence from the Literature and a Prospective Study," *Journal on Quality Improvement*, Vol. 26, No. 11, 2000, pp. 644-653.

Glassman, Peter A., et al, "Differences in Clinical Decision Making Between Internists and Cardiologists," *Arch Intern Med*, Vol. 157, 1997, pp. 506–512.

Glassman, Peter A., et al, "Physicians Personal Malpractice Experiences Are Not Related to Defensive Clinical Practices," *J Health Polit Policy Law*, Vol. 21, 1996, pp. 219–241.

Grossman, Michael, "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, Vol. 80, No. 2, 1972, pp. 223–255.

Hays, Ron D., and Toshi Hayashi, "Beyond Internal Consistency Reliability: Rationale and User's Guide for Multitrait Analysis Program on the Microcomputer," *Behavior Research Methods, Instruments, & Computers*, Vol. 22, No. 2, 1990, pp. 167-175.

Hofer, Timothy P., et al, "The Unreliability of Individual Physician 'Report Cards' for Assessing the Costs and Quality of Care of a Chronic Disease," *JAMA*, Vol. 281, No. 22, 1999, pp. 2098-2105.

International Finance Corporation, *Just What the Doctor Ordered: Prescriptions for Private Health Care in Developing Countries*, Washington DC: World Bank Group, Spring 1999.

Jones, Thomas V., et al, "Written Case Simulations: Do they Predict Physicians' Behavior?," *J. Clin. Epidemiol.*, Vol. 43, No. 8, 1990, pp. 805-815.

Kalf, Annette J., and Donna Spruit-Metz, "Variation in Diagnoses: Influence of Specialists' Training on Selecting and Ranking Relevant Information in Geriatric Case Vignettes," *Soc. Sci. Med.*, Vol. 42, No. 5, 1996, pp. 705-712.

Keeler, Emmet B, et al, "Hospital Characteristics and Quality of Care," *JAMA*, Vol. 268, 1992, pp. 1709-1714.

Keller, S.D., "Use of Structural Equation Modeling to Test the Construct Validity of the SF-36 Health Survey in Ten Countries: Results from the IQOLA Project," *J. Clin Epidemiol*, Vol. 51, No. 11, 1998, pp. 1179-1188.

Klitgaard, Robert, *Data Analysis for Development*, New York: Oxford University Press, 1985.

Litwin, Mark S., et al, "The UCLA Prostate Cancer Index. Development, Reliability, and Validity of a Health-Related Quality of Life Measure," *Medical Care*, Vol. 36, No. 7, pp. 1002-1012.

Lohr, Kathleen N., *Medicare: A Strategy for Quality Assurance.* Washington D.C.: National Academy Press, 1990.

Luck, Jeffrey, "How Well Does Chart Abstraction Measure Quality? A Prospective Comparison of Standardized Patients with the Medical Record, *Am J Med*, Vol. 108, No. 8, 2000, pp. 642-649.

Mathiyazhagan, K., "Willingness to Pay for Rural Health Insurance Through Community Participation in India," *Int. J. Health Plann. Mgmt.*, Vol. 13, 1998, pp. 47-67.

McGlynn, Elizabeth A., "Quality Assessment of Reproductive Health Services," *The Western Journal of Medicine*, Vol. 163, No. 3 (supplement), 1995, pp. 19-27.

McGlynn, Elizabeth A., "Six Challenges in Measuring the Quality of Health Care," *Health Affairs*, Vol. 16, No. 3, 1997, pp. 7-21.

Miller ER, et al, "Quality of Life in HCV-infection: Lack of Association with ALT Levels," *Aust N Z J Public Health*, Vol. 25, No. 4, 2001, pp. 355-61.

Mwabu, G., "Health Care Decisions at the Household Level: Results on a Rural Health Survey in Kenya," *Social Science and Medicine*, Vol. 27, 1986, pp. 315-319.

Newhouse, Joseph P., "Medical Care Costs: How Much Welfare Loss?," *Journal of Economic Perspectives*, Vol. 6, No. 3, 1992, pp. 3-21.

Nordyke, Robert, and John W. Peabody, *Capitation Evaluation Program (CEP) Round 1: Description and Fieldwork*, Santa Monica, CA: RAND, DRU-2026-WB, 1999.

Nordyke, Robert, *Privatization of Health Care Provision in a Transition Economy: Lessons from the Republic of Macedonia*, Santa Monica, CA: RAND, RGSD-155, 2000.

O'Neill, D, et al, "Variability in Scoring the Halchinsky Ischaemic Score," *Age Ageing*, Vol. 14, 1995, pp. 242–246.

Omaswa, F., et al, "Introducing Quality Management into Primary Health Care Services in Uganda," *Bulletin of the WHO*, Vol. 75, No. 2, 1997, pp. 155-161.

Page, A.C., et al, "Measuring Mental Health Outcomes in a Private Psychiatric Clinic: Health of the Nation Outcome Scales and Medical Outcomes Short Form SF-36," *Aust N Z J Psychiatry*, Vol. 35, No. 3, 2001, pp. 377-381.

Panamerican Health Organization (PAHO), *Nuevas Modalidades de Organización de los Sistemas y Servicios de Salud en el Contexto de la Reforma Sectorial: La Atención Gerenciada*, Washington, D.C.: PAHO, HSP/SILOS-40, 1995.

Peabody, John W., et al, "Comparison of Vignettes, Standardized Patients, and Chart Abstraction: A Prospective Validation Study of 3 Methods for Measuring Quality," *JAMA*, Vol. 283, No. 13, 2000, pp. 1715-1722.

Peabody, John W., et al, "Quality of Care in Public and Private Primary Health Care Facilities: Structural Comparisons in Jamaica," *Bulletin of the PAHO*, Vol. 28, No. 2, 1994, pp. 122-141.

Peabody, John W., et al, *Policy and Health: Implications for Development in Asia*, Cambridge, UK: Cambridge University Press, 1999a.

Peabody, John W., et al, *Reforming Health Care in Macedonia: RAND Technical Assistance Reports from July 1996 to May 1997*, Santa Monica, CA: RAND, DRU- 1649-WB, 1997.

Peabody, John W., Robert Nordyke, and Ninez Ponce, *Capitation Evaluation Program (CEP) Early Analysis of Round One Data*, Santa Monica, CA: RAND, DRU-2049-WB, 1999b.

Peabody, John W., *Will Measuring the Quality of Antenatal Care tell Us How it Works?*, Santa Monica, CA: RAND, RGSD-112, 1995.

Pebley, Anne R., and Noreen Goldman, *Social Inequality and Children's Growth in Guatemala*, Labor and Population Program, Working Paper Series 94-15, Santa Monica, CA: RAND, 1994.

Pécoul, Bernard, et al, "Access to Essential Drugs in Poor Countries," *JAMA*, Vol. 281, No. 4, 1999, pp. 361-367.

Phelps, Charles, "Diffusion of Information in Medical Care," *The Journal of Economic Perspectives*, Volume 6, Issue 3, Summer 1992, pp. 23-42.

Pieters, H.M, et al, "Simulated Patients in Assessing Consultation Skills of Trainees in General Practice Vocational Training: A Validity Study," *Med Educ*, Vol. 28, 1994, pp. 226–233.

Rao, Vijayendra, "Diet, Mortality, and Life Expectancy: A Cross-National Analysis," *Journal of Population Economics*, 1989.

Rethans, J.J, and C.P Van Bowen, "Simulated Patients in General Practice" a different Look at the Consultation," *BMJ*, Vol. 294, 1987, pp. 809–812.

Robalino, David, et al, *Does Fiscal Decentralization Improve Health Outcomes? Evidence from a Cross-Country Analysis*, Washington DC: The World Bank, Working Paper # 2565, 2001.

Robalino, David, *Social Capital, Technology Diffusion, and Sustainable Growth in the Developing World*, Santa Monica, CA: RAND, RGSD-151, 2000.

Robertson O, and MJ Dunbar, "Patient satisfaction compared with general health and disease-specific questionnaires in knee arthroplasty patients," *J Arthroplasty*, Vol. 16, No. 4, 2001, pp. 476-82

Roos, N.P, and L.L. Roos, "Small Area Variations, Practice Style, and Quality of Care," in *Why are Some People Healthy and Others Not? The Determinants of Health in Populations*, Evans, Robert G., et al, eds., New York: Aldine de Gruyter, 1994, pp. 231-252.

Rosenzweig, Mark R., and T. Paul Schultz, "Child Mortality and Fertility in Colombia: Individual and Community Effects," *Health Education and Policy Education*, Vol. 2, No. 3-4, 1982, pp. 305-348.

Rosenzweig, Mark R., and T. Paul Schultz, "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight," *Journal of Political Economy*, Vol. 91, No. 5, 1983, pp. 723-746.

Rostow, W.W., "The Take-Off Into Self-Sustained Growth," *The Economic Journal*, Vol. 66, No. 261, 1956, pp. 25-48.

SAS Institute Inc., *SAS/STAT Software: Release 6.12*, Cary, NC: SAS Institute Inc., 1996.

Seidman, Joshua, and Sheldon Greenfield, "Quality for Good Measure," *Seminars in Medical Practice*, Vol. 1, No. 1, 1998, pp. 7-10.

Shawyer, Richard J., et al, "The Role of Clinical Vignettes in Rapid Ethnographic Research: A Folk Taxonomy of Diarrhoea in Thailand," *Soc. Sci. Med.*, Vol. 42, No. 1, 1996, pp. 111-123.

Shuster, Mark A., et al, "How Good is the Quality of Health Care in the United States," *The Milbank Quarterly*, Vol. 76, No. 4, 1998, pp. 517-563.

Siu, Albert L., et al, "Choosing Quality of Care Measures Based on the Expected Impact of Improved Care on Health," *Health Services Research*, Vol. 27, No. 5, pp. 619-650.

Sriram T.G, et al, "Development of Case Vignettes to Assess the Mental Health Training of Primary Care Medical Officers," *Acta Psychiatr. Scand.*, Vol. 82, 1990, pp. 174–177.

StataCorp, *Stata Statistical Software: Release 7.0*, College Station, TX: Stata Corporation, 2001.

State of Victoria (SOV), *The Victorian Burden of Disease Study: Morbidity*, Melbourne, Victoria: Public Health Division, Department of Human Services, 1999.

Stocker, K., et. al., "The Exportation of Managed Care to Latin America," *NEJM*, Vol. 340, No. 14, April 8, 1999.

Thaver, Inayat H., et al, "Private Practitioners in the Slums of Karachi: What Quality of Care do they Offer?," *Soc. Sci. Med.*, Vol. 46, No. 11, 1998, pp. 1441-1449.

Vardi, Yehuda, "Nonparametric Estimation in the Presence of Length Bias," *The Annals of Statistics*, Vol. 13, No. 1, 1985, pp. 178–203.

Wasson, John H., et al, "A Randomized Trial of the Use of Patient Self-Assessment Data to Improve Community Practices," *Effective Clinical Practice*, Vol. 2, No. 1, 1999, pp. 1-10.

Weiner, Jonathan P., et al, "Variations in Office-Based Quality," *JAMA*, Vol. 273, No. 19, 1995, pp.1503–1508.

Weiss, M.G., "Cultural Models of Diarrheal Illness: Conceptual Framework and Review," *Social Science & Medicine*, Vol. 27, 1988, pp. 5-16.

World Bank (WB), *World Development Report 1993: Investing in Health*, Washington D.C.: WB, 1993.

Wouters, A., "Health Care Utilization Patterns in Developing Countries," *Bulletin of the WHO*, Vol. 70, No. 3, 1992, pp.381–389.

Ziv, Amitai, et al, "Lessons Learned form Six Years of International Administrations of the ECFMG's SP-based Clinical Skills Assessment," *Academic Medicine*, Vol. 73, No. 1, January 1998, pp. 84-91.